THE INDO-PORTUGUESE
LANGUAGE OF DIU

This dissertation is typeset using LaTeX.

THE INDO-PORTUGUESE LANGUAGE OF DIU

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties
ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op vrijdag 5 juni 2009, te 10:00 uur

door

Hugo Canelas Cardoso

geboren te Coimbra, Portugal

Promotor: Prof. dr. P.C. Hengeveld
Copromotor: Dr. U. Ansaldo

Overige Leden: Prof. dr. A. Baxter
                                   Prof. dr. J.C. Clements
                                   Prof. dr. J. Holm
                                   Prof. dr. J. Quer
                                   Prof. dr. M.I. Tomás
                                   Dr. N.S.H. Smith

Faculteit der Geesteswetenschappen

Contents

Acknowledgments

It would not be an exaggeration to say this research project was a collective effort, for even though the writing is my own it would have had no substance or actuality without the collaboration of many people. Listing their names by no means does justice to their generosity; I hope I will be able to make my appreciation clear to each and every one in other ways and on different occasions.

I am especially indebted to the speakers of Diu Indo-Portuguese who took their time to talk to me, who suffered my constant buzzing of the bell and tiresome quizzing. For their warm welcome, gracious cooperation and sincere friendship, I am deeply grateful to many Diuese, in particular the Cruz, Almeida, and Fernandes families - the most wonderful hosts and neighbours I could have wished for -, as well as the Abdul-Latif, Bica, Brito, Colimão, d'Souza, Duarte, Gomes de Mello, Ismail, Jessuclal, Kanji, Meggi, Natu, Rebello, Remédios, Rodrigues, Rosário and Tashi families. Among my most generous collaborators were the Christian schoolchildren, who certainly deserve special mention: thank you, Alexia, Alister, Armando, Ashley, Astle, Balfina, Conchita, Fabian, Jacob, Leslie, Liza, Neville, Ryan, Steven and all the others. I am also indebted to Fr. Ronnie d'Souza for his precious help in setting up my initial recording sessions and to Fr. Sérgio Mascarenhas for his interest and companionship.

Up there in the spotlight is also my family. Mãe, pai e Gonçalo, obrigado por tudo e mais alguma coisa, incluindo o vosso apoio e a vossa confiança. Eu sei que estas coisas não têm de se agradecer, mas vocês merecem a menção honrosíssima.

Back in Amsterdam, I could always rely on the enthusiastic and unflinching support of my supervisors Umberto Ansaldo and Kees Hengeveld, for which I am very grateful indeed. Jacques Arends, who is sadly no longer among us, was an essential driving force of the project in its early stages and one of my original supervisors; Jacques is sorely missed, and to him I dedicate the result of my research. The members of the Linguistics Department and the ACLC (both fellow doctoral students and teaching staff) as a whole have my recognition for turning Amsterdam into such a stimulating scientific environment. A host of colleagues have made essential and very different contributions to the success of this research project; for various reasons - ranging from bureaucratic help to undying friendship, including scholarly input, house-hunting support and general *gezelligheid* -, I am particularly grateful to Alexis Dimitriadis, Anne Baker, Antje Orgassa, Diana Apoussidou,

Ekaterina Bobyleva, Elly van den Berge, Els Verheugd-Daatzelaar, Enoch Aboh, Eva van Lier, Fred Weerman, Gerdien Kerssies, Gerry Wanders, Hans den Besten, Herby Glaude, Irene Jacobi, Jan de Jong, Jorge Gómez-Rendón, Lisa Lim, Maren Pannemann, Margot van den Berg, Margreet Dorleijn, Marije Michel, Marjolein Cremer, Niels Smit, Norval Smith, Olga Fischer, Rachel Selbach, Rafael Fischer, Robert Cirillo, Robert Cloutier, Roland Pfau, Sebastian Nordhoff, Victoria Nyst and Wieneke Wesseling.

Elsewhere in the world of Linguistics, I have been fortunate enough to receive all sorts of help and encouragement from Adrienne Bruyn, Alan Baxter, Arvind Bhandari, Clancy Clements, Dominika Swolkien, Donald Winford, E. Annamalai, Eric Anchimbe, Himanshu Upadhyaya, John Holm, Lachlan Mackenzie, Liliana Inverno, Maria Isabel Tomás, Nilotpala Gandhi, Robert Maier, Salikoko Mufwene, and many others whom I regularly met at conferences and symposia. I am deeply grateful for the stimulus they all provided.

My many and very good friends in Amsterdam - and you know exactly who you are - must also be thanked, and profusely so. I do not underestimate the importance of forgetting about research every now and then in order to be able to complete it; thank you all for being such a happy bunch, thank you for nurturing me (in more ways than one) and helping me unwind. And occasionally even helped me directly with my work. You have been more important than you probably realise.

List of abbreviations

1 - 1st person
2 - 2nd person
3 - 3rd person
A - Agent-like argument of transitive verb
ABL - Ablative
Adj - Adjective
AdjP - Adjective Phrase
Adv - Adverb
AdvP - Adverb Phrase
AG - Agent
ART - Article
AUX - Auxiliary
CAUS - Causative
CLS - Conjugational class
CMP - Complementiser
COLL - Collectiviser
COM - Comitative
COMP - Comparative marker
COND - Conditional
COP$_i$ - Individual-level copula
COP$_s$ - Stague-level copula
DAT - Dative
Dct - Deictic modifiers
DEM$_d$ - Distal demonstrative
DEM$_p$ - Proximal demonstrative
DIM - Diminutive
DIP - Diu Indo-Portuguese
EMPH - Emphatic
Eng - English
EXCL - Exclusive
EXS - Existential
f - Feminine
FUT - Future
Guj - Gujarati

HORT - Hortative
INCL - Inclusive
IND - Indicative
INF - Infinitive
INS - Instrumental
Int - Intensifier
INTJ - Interjection
IP - Indo-Portuguese
IPA - International Phonetic Alphabet
IPFV - Imperfective
IRR - Irrealis marker
Konk - Konkani
LOC - Locative
m - Masculine
n - Neuter
N - Noun
NEG$_{cl}$ - Clausal negator
NEG$_{ct}$ - Constituent negator
NEG$_p$ - propositional negator
NP - Noun Phrase
NPST - Non-Past
OBJ - Object
OBL - Oblique
Ord - Ordinals
p - Plural
P - Patient-like argument of monotransitive verb
PASS - Passive
PFV - Perfective
PL - Plural
PoS - Parts-of-Speech
POSS - Possessive
PP - Prepositional Phrase
PREP - Preposition
PROG - Progressive

PRS - Present
PST - Past
PTCP - Participle
Ptg - Portuguese
PURP - Purposive subordinator
Qtf - Quantifier
R - Recipient argument of ditransitive verb
Rel - Relative clause
REL - Relativiser
REQ - Requestative
S - Single argument of intransitive verb

SBJ - Subject
SBJV - Subjunctive
SG - Singular
SIML - Similative plural marker
SP - Standard Portuguese
SVC - Serial verb construction
TAM - Tense, Aspect and Mood
U - Undergoer argument of ditransitive verb
V - Verb
VP - Verb Phrase

Dedicated to the
memory of
Jacques Arends

Part I

Background

Chapter 1

Introduction

The aim of this dissertation is to provide an accurate description of Diu Indo-Portuguese henceforth 'DIP', a contact language of India with a history spanning nearly five centuries. The description is not only linguistic but also - given the interaction between sociohistorical conditions and the output of language contact[1] - historical and sociodemographic. In this introductory chapter, I will begin by locating Diu Indo-Portuguese in space and within the larger context of the Indo-Portuguese (creole) languages, before setting out the structure of the dissertation.

1.1   Diu

Diu is an island located off the coast of Gujarat, at the southern tip of the Saurashtra peninsula (formerly know as the Kathiawar region) in India - see Map 1.1.

The island of Diu is nearly 15 kms long and no more than 5 kms wide, and it is separated from the mainland by a tidal creek which is presently spanned by two bridges. As an administrative unit (corresponding to the territory controlled by Portugal between the 16th century and 1961), Diu includes not just the island proper but also a short stretch of the mainland where the village of Goghla is located, and the small enclave of Simbor some 40 kms away from the island.

The creek opens up into a well protected harbour in front of Diu Town, the urban centre of the territory located at the eastern tip of the island; Diu Town nowadays concentrates nearly half the population of the entire territory - updated demographics can be found in section 2.1. The remainder of the island is dotted with villages, the most important of which are Fudam (known to the Portuguese as 'Podamo' or ´Fodão'), Bucharwada (formerly 'Buxinvará'), Vanakbara (formerly 'Brancavará') and Malala. Goghla, across

---

[1] A strong version of the social determinism in language contact is articulated by Thomason and Kaufman (1988:35):

> [I]t is the sociolinguistic history of the speakers, and not the structure of their language, that is the primary determinant of the linguistic outcome of language contact.

Map 1.1: Map of India

the water, is also an important village. Map 1.2 is a 1953 chart of the territory designed by C. Ferrão for the government of the *Estado da Índia*.[2]

After the decolonisation of Diu, Daman and Goa - the three longest-standing Portuguese possessions in India - in 1961, the three discrete territories plus Dadra and Nagar-Haveli (see chapter 3, Map 3.1) were grouped under a Union Territory, i.e., an administrative unit directly dependent on the central government. Goa achieved statehood in 1987, and nowadays the union territory is circumscribed to Daman, Diu, Dadra and Nagar-

---

[2] The designation of the political unit comprising the Portuguese possessions of India, though originally 'the designation of «Portuguese India» had a wider meaning during the 16th, 17th and 18th centuries, and included not just the territories of Goa, Daman and Diu but likewise the cities, factories and fortresses established from the western coast of Africa to the Middle East and also on the Indian continent, Ceylon, Malaysia, Moluccas and the several religious missions, including those of Japan, and still other commercial posts along the vast maritime trajectory described by the Portuguese. With its capital in Goa, the viceroys and governors had jurisdiction over those territories, which included Mozambique until 1752 and Macau, Solor and Timor until 1844.' (my translation of Morais 1997:9).
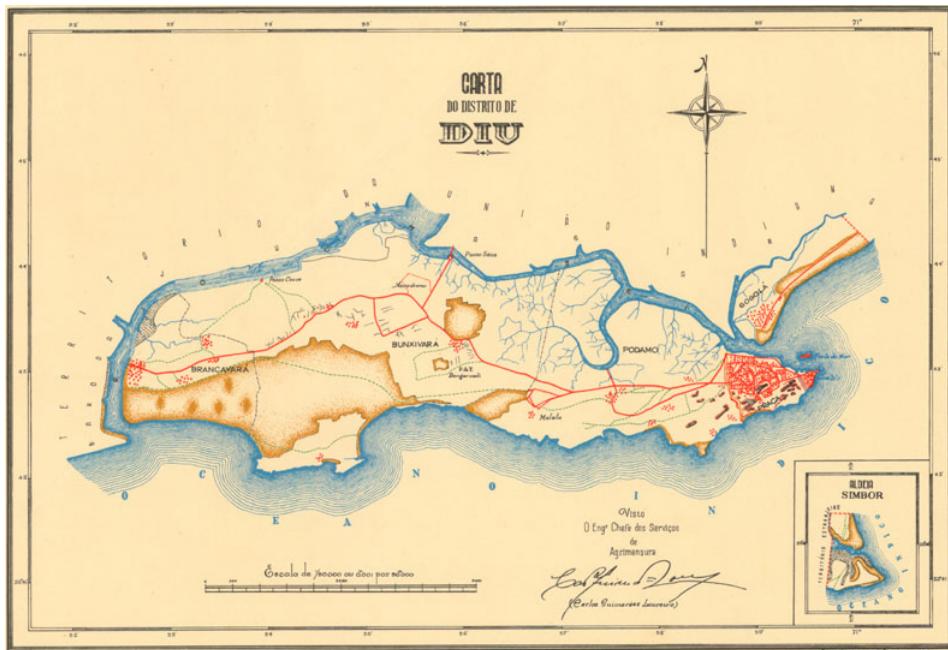
Map 1.2: Colonial map of Diu island and dependencies

Haveli, governed from Daman.

1.2   Indo-Portuguese

The label 'Indo-Portuguese' (henceforth IP) does not stand for one language but rather a number of Portuguese-lexified contact languages scattered across South Asia, which resulted from the role of the Portuguese language in the region and the colonial activities of the Portuguese from the 16th century onwards. Languages identified as Portuguese-based creoles were once spoken on several coastal areas of India and Ceylon, as indeed across Asia; a hypothesis for the genesis and spread of restructured Portuguese in Asia and within the Indian subcontinent relates to an alleged Portuguese-based pidgin which served as a trade language and even intercultural medium of communication from East Africa to Japan.[3] In South Asia, relatively stable varieties of IP were at some point to be found in Sri Lanka, in several ports of present-day Gujarat (e.g. Diu, Daman, Bassein, Surat) and Maharash-

---

[3] Portuguese functioned as a *lingua franca* in this region for centuries, and here other Portuguese-lexified creoles also formed: e.g. Malacca (see Baxter 1988), Macau (see Vasconcellos 1892, Batalha 1959), Batavia/Jakarta and Tugu (see Schuchardt 1890), and East Timor (see Baxter 1990). For a collection of 16th- through 19th-century documents indicative of the role of the Portuguese language in Asia, see Lopes (1936) and Matos (1968).

tra (e.g. Bombay, Chaul/Korlai, Thane), Goa[4], as well as Karnataka and Kerala (e.g. Mangalore, Cannanore, Tellicherry, Mahé, Calicut, Cochin, Quilon) but also on India's Eastern coast (e.g. Meliapor, Nagappattinam, Tranquebar, Pondicherry, Pipli, Calcutta) and Bangladesh (e.g. Dhaka, Chittagong); the varieties of Indo-Portuguese recorded are shown on Map 1.3.
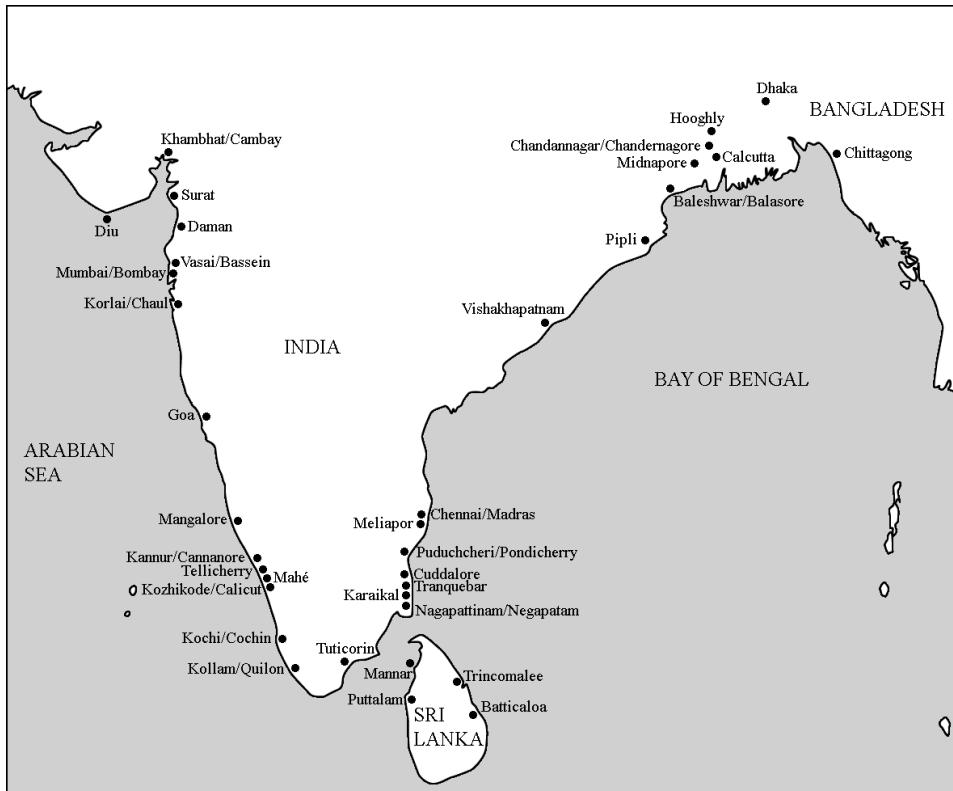
Map 1.3: Recorded *loci* of (Indo-)Portuguese in South Asia

Despite the observed similarities between them (Ferraz 1987, Clements 2000), the IP languages differ substantially from each other. The Indo-Portuguese family can be subdivided into clusters according to the endemic language of its physical environment and ultimately - because the different varieties are geographically discrete and significantly divergent - classified as individual languages. Schuchardt (1899) proposed a classification of the Asian Portuguese-based creoles according to the typology of their substrate; IP was

---

[4]The historical existence of IP in Goa is the matter of some debate. Analogy with the other strongholds may suggest so (cf. Clements 2000) and Dalgado (1921) assumes it, but there is no clear attestation of the fact; Goa's position as the capital of the *Estado da Índia* may have meant a stronger pressure of standard Portuguese, to the effect of blocking the formation of a greatly restructured variety.

therefore divided into *Gauro-Portuguese* (those with Indo-Aryan substrate, viz. Gujarati and Marathi) and the *Dravido-Portuguese* varieties (those with a Dravidian substrate, such as Mallayalam). Going into even more detail, the cluster of IP varieties commonly known as *Norteiro* 'Northern' (see Dalgado 1906), which falls squarely into the proposed Gauro-Portuguese category, includes the varieties spoken in the former *Província do Norte* 'Northern Province', the stretch of land between Daman and Chaul/Korlai - including the Bassein/Bombay area - and usually meant to encompass the island of Diu.

As a whole, and considering that systematic contact between South Asian languages and Portuguese began as early as 1498, IP has a history of over five centuries.[5] At present, Indo-Portuguese varieties are spoken with some degree of vitality in Diu (see also Cardoso 2006b), Daman (Clements and Koontz-Garboden 2002) and Korlai (Clements 1996). Two moribund varieties subsist in Cannanore, present-day Kannur, and Cochin, present-day Kochi.[6] The vitality of the variety of Sri Lanka recorded in the 1980s (Smith 1977, 1984, Jackson 1990) cannot be ascertained at present due to ongoing conflict, but it is likely that the language is still actively spoken.

1.3   Diu Indo-Portuguese

Diu Indo-Portuguese refers to the variety of Indo-Portuguese spoken in the territory of Diu. The designation is strictly scientific, because on the island no taxonomic distinction is recognised between DIP and standard Portuguese (henceforth SP). This does not mean the inhabitants of Diu fail to apprehend the differences. It may be so in the case of those with no knowledge of any of the two codes, including the political authorities, but (native) speakers are very much aware of the distinction - see section 2.2.1.2 for an expanded description of this issue. Both DIP and SP are therefore subsumed under the designation of *Portuguese*. The native speakers themselves use a range of expressions when it becomes necessary to distinguish the local variety from SP: *purtəgez də diw* 'Portuguese of Diu', but also *lĩg tɔrt* 'twisted language', *purtəgez kebrad* 'broken Portuguese', *purtəgez barat* 'cheap Portuguese' or *lĩg də trap* 'ragged tongue'. The depreciative nature of most of these epithets is a powerful indication that, among DIP-speakers, DIP does not command much overt prestige relative to SP (see section 2.2.1). I have therefore opted for the term *Indo-Portuguese*, for various reasons:

- *Indo-Portuguese* has for a long time been used as a cover term for the various Portuguese-lexified contact languages of South Asia (see section 1.2), so the treatment of the Diuese variety as *Diu Indo-Portuguese* has the virtue of highlighting its membership

---

[5] In this period, however, it has met with different fates in different locations. Most of the varieties have since died out, whereas in those places where IP still survives its degree of stability or endangerment varies quite considerably. Early sources which take a broad look at the Indo-Portuguese varieties as a cluster include Teza (1872), Coelho (1880), Schuchardt (1899) and Vasconcellos (1901). More recent accounts include Theban (1985), Clements (1991), Tomás (1992) and Cardoso (2006a).

[6] The survival of the varieties of Cochin and Cannanore, which had been presumed extinct (see e.g. Smith 1995), was confirmed during field visits to these two locations, in 2006 and 2007. I have been able to encounter 6 speakers of Cannanore IP and only 1 speaker of Cochin IP - this finding was briefly reported in Cardoso (2006a).

of this cluster, which share not only part of their formative history but also several typological traits (see e.g. Ferraz 1987);

- the term *Indo-Portuguese* has become current as a designation of clearly-defined products of the interaction and contact between Portuguese and South Asian cultures, on a variety of levels (e.g. architectural and decorative styles, cuisine, genealogy, music, religious practices, etc.); I find it advantageous to stress that language admixture is one among several manifestations of cultural encounter - though not necessarily formed by the same processes (cf. Palmié 2006) -, which is achieved with the selection of *Indo-Portuguese* as a label for the language;

- the genealogical duality of DIP from a phylogenetic perspective is also conveniently captured by the expression *Indo-Portuguese*, which implies the contribution of both an Indic element and a Portuguese element. The contribution of the various languages involved in the formation and development of DIP is clarified in chapter 9.

The absence of the term *Creole* from the designation adopted is purposive. I base my decision on three factors:

- the high specificity of the linguistic notion of *creole* is not transparent to non-specialised observers, and often ideologically charged even to those within the field, as denounced in the debate of 'creole exceptionalism' (see e.g. DeGraff 2003);

- a word *kriol* does exist in DIP, referring to a child from an underprivileged background raised as a servant for a wealthier family.[7] This fact enhances the potential for misinterpretation of the label *Creole* among the very speakers of the language in question;

- the academic field of Creole Studies has now recognised the fuzziness of the terminological boundaries of the concept and the multiplicity of criteria which have shaped various competing definitions of *creole* (see e.g. Jourdan 1991, Baptista 2005). According to a strictly sociohistorical definition of a creole language which would bring to the fore the similarities between the formative contexts of many of these contact varieties (i.e., the contact originated by efforts of European expansion), DIP would be easily classified alongside creoles. On the other hand, narrow formal characterisations of a typological class of 'creole languages', such as that advocated in McWhorter (1998), would probably exclude DIP from its ranks.[8] By avoiding the overt classification of DIP as *creole* I am not denying the relevance of this research to the field of Creole Studies (and vice-versa) but simply avoiding a terminological complication which is often purely theory-internal and not necessarily beneficial to theory itself.

Whereas certain contact languages have an early body of written records, no matter

---

[7] The same word, spelt *criolo* but with a similar meaning, is reported for Goan Portuguese in Dalgado (1900).

[8] One of the least prototypically creole traits of DIP, in light of McWhorter's 'Creole Prototype', is the significant use of verbal inflection. For further information, see section 6.2.

how scarce[9], other contact varieties have not been documented during their early stages. In the case of DIP, the first systematic data collection was that organised by the governor Pedro Francisco Perry da Câmara in 1882 at the request of Hugo Schuchardt[10]; having reached Schuchardt, the short corpus resulted in the publication of Schuchardt (1883), the only linguistic description of the language until the present one and a precious source of data concerning late 19th-century DIP.

The publication of Schuchardt (1883) deserved an acknowledgment on the part of the Goan philologist Sebastião Rodolpho Dalgado (Dalgado 1903) and, most crucially, a somewhat detailed comment by an educated local resident, Jeronymo Quadros[11], between 1902 and 1905 (Quadros 1907). For a long time after these documents, which attested to some vitality of the language, little else was written on DIP, leading to the general assumption of its extinction. In the 1990s, having visited the island, both (Clements 1991) and Tomás (1992) report on the continuing use of DIP, even though they could only provide a brief description of the linguistic community.

Indo-Portuguese (and DIP in particular) was central to the earliest inquiries into creole language formation, as attested to by the interest of Schuchardt, Dalgado and their contemporaries, but these varieties (as well as the contact languages of Asia and the Pacific in general) have received relatively little attention since. Parkvall, for instance, recently noticed that the overwhelming majority of the articles published in the *Journal of Pidgin and Creole Languages* is dedicated to the languages of Haiti and Suriname, while 'important contact languages such as the Gulf of Guinea Portuguese Creoles, Indo-Portuguese, the Creoles of the Philippines, the Russo-Siberian pidgins (. . .) remain virtually untouched' (Parkvall 2002:360). The present research intends to contribute to solving the inbalance, so as to enhance the accuracy of theories of language contact. Another important motivation for undertaking this project was the realisation of the language's possible endangerment, which is described in section 2.4.

1.4   Structure

This dissertation is organised into three parts, labelled 'Background', 'Description' and 'Discussion' respectively. I will briefly describe the contents of each of these parts, in turn.

Part I, 'Background', provides the non-linguistic information necessary to understand the formation and development of DIP. Chapter 2 is a synchronic account of the setting

---

[9]See, for instance, Berrenger (1811) and the Hugh Nevill Manuscript (Jackson 1990, Jayasuriya 1996) for Sri Lanka Indo-Portuguese; Schuchardt (1890) for Batavia/Tugu Creole Portuguese; Arends and Perl (1995) for the Surinamese Creoles in general; van den Berg and Arends (2004) for Sranan.

[10]The original manuscripts are archived at the *Hugo Schuchardt Nachlaß*, at the University of Graz, with the references 8752 through 8760.

[11]According to Matos (1999:9), Jeronymo Quadros was a native Diuese who worked there as a primary-school teacher before taking up administrative positions in the political structure of the *Estado da Índia*. He had a keen interest in the history and ethnography of Diu, which resulted in his publication of the 16th-century *Tombo de Diu* (*Goa Historical Archives*, doc. 624 - 'Tombo de Diu'), a volume of historical sources for the history of the Portuguese in Diu 1citepQuadros1899 and, in the early 20th century, various contributions to a Portuguese newspaper concerning the circumstance of Diu (collated in Quadros 1907). Jeronymo Quadros died in 1947.

in which DIP interaction takes place. It offers information on the composition of the population of Diu, with particular focus on the Roman Catholic community (which, as established in 2.1, constitutes the native speaker community of DIP), as well as on patterns of multilingualism among the speakers of DIP. This chapter further includes a brief study of the variation observed in DIP, and also an assessment of the language's present and future vitality. Chapter 3 charts the social history of Diu in order to account for the formation and diffusion of DIP. It begins by presenting a brief political and social history of the territory, and the Portuguese possessions in India as a whole, and finally proposes a scenario for the formation of the Diuese variety of IP.

Part II contains the synchronic linguistic description of DIP. Chapter 4 provides the necessary descriptive preliminaries, such as the methods of data collection and treatment, the composition of the corpus, as well as the relevant conventions of phonemic/phonetic transcription, orthography and glossing. The linguistic description proper is provided in chapters 5 through 8. Chapter 5 describes the phonology of DIP, and chapter 6 defines the Parts-of-Speech (lexical categories) relevant for the language while simultaneously describing the morphological operations (with the exception of derivational morphology) available to DIP. Chapter 7 is an overview of the language's syntax focusing on its most salient features and also the interface between syntax and the expression of pragmatic values. Finally, chapter 8 is a broad description of the lexicon of DIP - including derivational morphology -, and it touches on issues which are particularly relevant for a comparative study and for the correct definition of the language's formation.

Part III, 'Discussion', consists of a comparative study which builds on the data provided in Part II, as well as information concerning the languages most relevant for the formation of DIP. The intention is to ascertain the varied provenance of particular characteristics of the language, thereby illustrating the complexity of the process of genesis/development of DIP.

Chapter 2

Diu Indo-Portuguese at present

This chapter draws a picture of the present-day use of DIP and SP within the territory of Diu. When dealing with statistics of language use, it is inevitable to speak of DIP and SP together for a number of reasons:

a) Official statistics and censuses do not distinguish the two registers;

b) It is not uncommon for speakers of DIP to have some knowledge of SP;

c) Prestige asymmetries between the two registers and some specific sociolinguistic dynamics among the speech community mean that precise quantification of DIP speakers vis-à-vis SP speakers is virtually unattainable.

I will therefore begin by describing the distribution of knowledge of *Portuguese* (as a cover term for both the standard variety and the local variety), attempting to fine-tune the distribution of one or the other registers whenever possible. Next, I will explore the ways in which speakers of DIP relate to other languages present in the territory and/or historically relevant, including standard Portuguese, Gujarati, Daman Indo-Portuguese, Hindi and English. Finally, I will describe to what extent both the distribution of DIP/SP across the Diuese population and the relative prestige assignment are products of the island's recent past.

2.1   Demographics

All across modern India, religion is one of the most fundamental ethnological variables, which determines several social manifestations (e.g. rituals, social relations, marriage, language, dress, cuisine), and Diu is not an exception. The major religious groups in Diu are - in descending numerical order - Hindus, Muslims and Christians (Roman Catholics), and as it turns out the various communities relate differently to what is widely interpreted as the language of the former colonists. In addition, census information - in particular of the colonial period - is very often organised along religious lines, which makes it advantageous to break up this demographic section accordingly.

I shall begin with the Catholic section of the population for the very specific reason that the Catholics are presently the only native speakers of DIP, and the population at large tends to associate *Portuguese* (both DIP and SP) with the local Christians.

2.1.1 Roman Catholics

The Christian community in Diu is entirely Roman Catholic, which is not surprising given the fact that the evangelisation of the island has been, since the 16th century, the prerogative of the staunchly Catholic Portuguese. The Portuguese were responsible for the introduction of Christianity in Diu as soon as they gained control over the fort (see 3.2.3) and, even today, the Christians are widely seen as their cultural descendants; among the Diuese, it is common for the Portuguese language to be called the 'language of the Christians'. While both DIP and SP are to be found among other communities, this epithet is partially justified. In fact, the Christians are, *grosso modu*, those for whom DIP is a first language; unlike any other community, the vast majority of Christians use DIP in everyday life and within a family context. If we approach the distribution of DIP from a generational perspective, we observe that the only children and youngsters who speak DIP at all are those of Christian descent.

During the colonial period, the Christians had particularly privileged access to administrative posts and Portuguese-medium education, with the consequence that they enjoyed relatively high social status at the time. This is not necessarily the case at present. Singh et al. (1994), referring indiscriminately to both Diuese and Damanese society, write that

> They [the Christians] enjoyed high social position during the Portuguese regime being placed high in administration. Presently, they perceive their position as inferior to the Brahman, Vania, Koli Patel, Kamli, Bhandari, Sagar and Bari, and superior to the Machhi, Mangela, Mitna, Dhodia, Dubla, Momin and Mahyavanshi [caste distinctions external to Christianity]. (Singh et al. 1994:51)

It is not clear whether ranking the Christians in relation to the local Hindu caste hierarchies is an entirely sound approach, but the point to be retained is that of a certain social decay following the 1961 integration in India. Nonetheless, one still has the impression that the Christians of Diu have a social visibility incommensurate with their numbers, which is probably a result not only of the jobs they mostly take up (administration, health-care and teaching) but also because they tend to cluster in a particular area of Diu Town[1] (pop. 21,576 in 2001). Furthermore, the various churches and public Christian celebrations are very conspicuous on the island, and there is a prestigious Christian school with branches in Diu Town, Vanakbara and Goghla. The fact that the historical significance of former Portuguese rule has to some extent been transferred onto the Christian population is probably not alien to the enhanced attention this small community enjoys in the territory.

---

[1] The Christians have since 1961 converged exclusively into Diu Town, whereas earlier there were Christian pockets in other villages of the territory, notably Fudam and Vanakbara; a small Christian population is reported for Vanakbara in Brito (1998), a survey of the territory made in the 1950s. This small exodus was reported by some of the oldest members of the Christian community, some of whom spent their early years outside Diu Town, and is evidenced by the various churches and chapels outside the town.

Official census data from 1991 and 2001 is not ideal in the sense that they refer to both Daman and Diu and the figures for the two territories are not always provided separately. The total population figures for the territory of Diu are 39,488 in 1991 and 44,110 in 2001. However, the same censuses fail to discriminate between the Damanese and Diuese figures concerning the division according to professed religion; in 1991, the Christian population of the two territories numbers 2,904 people, but the fact that this figure lumps together the communities of Diu and Daman makes it impossible to get a precise impression of the local Diuese picture. One is therefore left to rely on data provided by the parish priests, who report that at present about 250 Christians, divided over some 35 households, reside in Diu. Assuming the population of Diu has not changed significantly since 2001, the approximate percentage of Christian residents is 0.6%.

Whereas most Christians are native Diuese, Diu has also witnessed the more or less recent settlement of some Damanese and Goans - either spouses of local Diuese or part of entirely non-Diuese families - as well as some Catholics from other areas of India, mostly from southern India (Kerala and Tamil Nadu in particular) and the Mumbai region. Of the 250-odd Christians estimated in Diu at present, around 50 are originally from areas other than the former Portuguese colonies of Diu, Daman or Goa; they tend to have no knowledge of DIP or SP. Of the remaining 200, one must allow for a small margin of those who, for one reason or another, have little knowledge of any of these languages. Overall, it can safely be advanced that around 170 of the Diuese Catholics use DIP or SP, or both, as their everyday language(s).

2.1.2 Remaining population

According to the 1991 census, some 93% of the Diuese population was Hindu, whereas the Muslim community made up nearly 6% of a population of 39,488. The figures therefore approach 37,000 for the Hindus and 3,400 for the Muslims, in 1991. There is also a small Jain[2] community, and the last Diuese Parsis[3] are said to have left the territory only around 1950.

All throughout the colonial history of Diu, many non-Christians have used what is referred to in various sources as *Portuguese* which once again leaves us wondering what register is being referred to. Close to the end of the 19th-century, Quadros says the following concerning the role of Portuguese in Diu:

> O idioma de que usam os não christãos são o guzerathe e o mussulmano, havendo alguns que

---

[2] Jains are the followers of Jainism, a religion which originated in the Indian subcontinent around the 3rd century BCE. Although Jains reside all throughout India and abroad, it is the northwestern states of Gujarat, Maharashtra and Rajasthan which host the largest communities. It is not clear how many live in Diu at present, but the 1920 census counts a population of 142 (quoted in Delduque 1928).

[3] The Parsis descend from a group of Zoroastrians who migrated from Persia to Western India sometime between the 8th and 10th century AD. Although the *circa* 75,000 strong community is present in several areas of South Asia, there is a particularly high concentration around the city of Mumbai (Bombay) and the South Gujarat region centered around Surat. In 1920, according to the census (quoted in Delduque 1928), 74 Parsis resided in Diu. For the particular link of the Parsi community to the island of Diu, see section 3.2.3.

fallam, lêem e escrevem o portuguez. (Quadros 1899:97)

‘The language used by the non-christians are gujarati and muslim, some of them being able to speak, read and write portuguese.’

The reference to *mussulmano* is somewhat puzzling. The author might be referring to Urdu. At present, the native language of the overwhelming majority of the population (Muslim or Hindu) is Gujarati, and the non-Christians who speak DIP or SP do so non-natively. A common claim among middle-aged to senior Muslim citizens is that most of their generation are able to speak Portuguese, which is partly echoed by the corresponding Hindu generation although with less assertiveness. Meaningful though it is, this claim unfortunately does not clarify what degree of proficiency is being referred to, or whether the language mentioned is SP or DIP.

Muslims and Hindus, unlike Christians, live in the villages of the territory as well as Diu Town. Some of the speakers of Portuguese encountered were inhabitants of the villages, in particular Goghla, Vanakbara, Fudam, Bucharwada and Gandhipara (outside the city walls, see Map 1.2). The village of Fudam is noteworthy because a significant part of its population (overwhelmingly Hindu) has emigrated to Portugal. The Diuese who can prove that their families lived in Diu before 1961 are still entitled to a Portuguese passport, which, despite the chronic moroseness of the bureaucratic process, has led to massive migration to Portugal. Given that most of the island's population is Hindu, this is also the most significant parcel of the emigrant population (see Lourenço 2005). During Portuguese rule, mobility between the different colonies under Portuguese rule was relatively easy, and the *Estado da Índia* established a close link with Mozambique. Among the Hindus, migration to Portugal has recently replaced the previous pattern of migration to Mozambique. Those who have settled in Portugal do return regularly to Diu, in particular to participate in ceremonies such as weddings. It is certainly an interesting observation that, particularly in the months of December through March, the territory witnesses a significant influx of families who ally obvious wealth to (near-)native proficiency of standard Portuguese. The sociolinguistic implications of the fact are explored in section 2.2.1.2 below. The Muslim community also has a long history of emigration (mostly to Mozambique), often not for life. It is not uncommon for native Diuese Muslims to settle back in Diu after a period of emigration.

Prior to 1961, many Hindu and Muslim citizens also attended Portuguese-medium schools. In fact, among the older Hindus and Muslims who experienced the colonial rule, primary education is the reason most frequently invoked for their command of Portuguese. The prototypical employment of Muslim men in Diu is that of shopkeepers and, in this capacity, they reportedly interact in SP or DIP with some of their customers, in particular the Christians. This claim has been bourne out by observation: in the traditional market area of Diu Town, many older Muslim shopkeepers were conversant in Portuguese, and their fluency ranged from basic and formulaic to near-native.

This description of the distribution of language knowledge must remain impressionistic given that, during fieldwork, it was not possible to carry out a full survey of language use on the island. In the absence of official statistics concerning the use of Portuguese among the Hindu and Muslim communities, it is therefore very difficult to advance any concrete

figures.

2.2   Multilingualism

From a linguistic point of view, Diu is a highly diverse ecology - as are many communities across India (see e.g. Pandit 1979, Annamalai 2003, Saxena 2005) -, and DIP speakers relate in various ways to the various languages in their linguistic environment. Chief among these is Gujarati, the main language of the region and native language of the overwhelming majority of the Diuese population. In addition, Hindi and English have recently gained prominence in the territory, as they have in most of India. There is also a very close historical connection - political as well as cultural and religious - between the territories of Diu and Daman; population movements and intermarriage, among other things, have resulted in close linguistic contact between the Indo-Portuguese variety of Diu and that of Daman, which at present are actively used in each other's traditional territories. Finally, there is a long - and crucial - history of SP being used in Diu, which remains to this day.

Approaching language use in a multilingual space begs the concession that linguistic behaviour (including linguistic allegiance and purism) is instrumental in defining, blurring or manipulating identity and social structures (e.g. Foley 2005, Saxena 2005). Labov (1972:111) recognises the usefulness of language production as a 'sensitive index of many [. . .] social processes', and one which is malleable to some extent. He observes

> Variation in linguistic behavior does not in itself exert a powerful influence on social development, nor does it affect drastically the life chances of the individual; on the contrary, the shape of linguistic behavior changes rapidly as the speaker's social position changes. (Labov 1972:111)

The notion that linguistic behaviour accompanies social change is crucial when approaching social settings which have witnessed a sudden redefinition of power, such as postcolonial societies.

2.2.1   Diu Indo-Portuguese and Standard Portuguese

I will begin the discussion of multilingualism among DIP speakers with the linguistic equation which is arguably the most complex, viz. that pertaining to DIP and SP. Part of the complexity of this question stems from the fact that, sociolinguistically, the two are often subsumed under a single category. As a result, SP is allowed to retain a certain normative power even among speakers with little or no contact with it. When combined with historical qualitative notions of asymmetric prestige, this brings about some interesting sociolinguistic patterns which are explored in section 2.2.1.2. Section 2.2.1.1 accounts for the synchronic role of SP in Diu.

2.2.1.1 Modern sources of SP

We have established before that the speakers of DIP assign their language relatively low overt prestige (cf. the language's depreciative epithets reported in section 1.3) and also that standard Portuguese is held as a norm. It is common for those who, by virtue of family tradition or education, have knowledge of SP to regard their own use of DIP as a communicative concession, and there is a general wish for further access to SP; in this equation, DIP is widely regarded as 'incorrect Portuguese'.

The present-day relevance of SP in Diu has a number of different causes. Despite the transfer of sovereignty in 1961, the cultural ties between Diu and Portugal have not been entirely suppressed, as evidenced by the possibility for many Diuese to obtain a Portuguese passport (see section 2.1 above). Both the more recent wave of migration to Portugal and the longer-standing connection with Mozambique (where many Diuese families are still represented) are effective sources of contact with SP. So is, to a lesser extent, migration to Goa, where a small portion of the (Catholic) population still speaks a variety of Portuguese very close to the European standard. For the past few years, Diu has also received a Portuguese television channel through cable.

Additionally, the Catholic priests operating in Diu are mostly Goa-educated. Until recently, all priests were fluent in SP and conducted mass in this language. Lately, only some of the priests assigned to the parish of Diu speak Portuguese, and even though most ceremonies are now conducted in English, religious service is still delivered in Portuguese once a week at least. Besides, in the recent past it was not uncommon for Goan civil servants and military to be stationed in Diu, so that there was a constant contact with Goan Portuguese, which has the peculiarity of having borrowed various words from Konkani. It is therefore not surprising that the DIP lexicon contains a significant Konkani element (see section 8.7.2); Goan Portuguese, rather than European Portuguese, may well be the actual and most accessible standard recognised by speakers of DIP.

It may be striking that, in a postcolonial society, the language of the former rulers should still play such an active role. The following section will try to clarify that not just the role of SP but also the connotations of this language are an ongoing reflection of a sociolinguistic situation set up during the colonial rule.

2.2.1.2 Language maintenance in a postcolonial society

Bourdieu (1982, 1991) approaches linguistic practice in society from the perspective of an economic metaphor in which different languages are assigned different relative *values*, expressed in terms of putative correctness or prestige. In his model, the processes of value assignment are socially and politically motivated and depend on notions of power asymmetry. These considerations are particularly relevant for postcolonial societies, given that the moment of decolonisation (by whatever means) normally involves, if not a reversal, at least a redefinition of power. In the case of Diu, and with regard to the fate of the Portuguese language, decolonisation resulted in complete lack of official support for the maintenance of the language of the former colonists. It is not an exaggeration to state that, given the general lack of distinction between SP and DIP, the local variety has suffered by proxy: the anti-colonial sentiment of certain sections of Indian society is tendentially

adverse to any reminder of what is considered a period of European occupation. In such a scenario, long-term maintenance of both DIP and SP was only possible because they commanded strong allegiance.

On the part of the native speakers of DIP, who are most aware of the distinction between their language and SP, allegiance to one and the other probably works differently. Historically, the relative prestige of DIP and SP have been fundamentally different, with SP clearly taking the upper hand during the colonial period as a) the language of education, and in particular of higher education in Goa or Portugal, b) the language of the emigrants, c) the language of the Church authorities and, crucially, d) the language of the ruling elite and those with close contact therewith. Cultured people in the *Estado da Índia* seem to have had some contempt for DIP. This is patent in the following opinion expressed by Jeronymo Quadros, a learned man acquainted with the Diuese reality and otherwise relatively supportive of Schuchardt's (1883) documentation effort:

> Os usos e costumes dos christãos de Diu são os mesmos ou quasi mesmos, que os dos christãos de Goa e Damão. Fallam o portuguez, mais ou menos correctamente, não sendo todavia raros certos idiotismos, solecismos e barbarismos intoleraveis. (Quadros 1899:98)

> 'The habits and traditions of the christians of Diu are the same or nearly the same as those of the christians in Goa or Daman. They speak Portuguese, more or less correctly, though certain intolerable idiocies, solecisms and barbarisms are common.'

The situation seems to have been comparable elsewhere in the Indo-Portuguese-speaking regions. Concerning the dialects of the Bombay area in 1906, Dalgado is quite explicit about this issue:

> As classes ilustradas manifestam desamor à sua língua materna, pela consciência e pejo que têm da sua corrupção, e procuram descartar-se dela, servindo-se ou do português legítimo ou do inglês, língua oficial, principalmente em Bombaim e nos subúrbios. (Dalgado 1906)

> 'The educated classes reveal lack of love for their mother tongue, because of their conscience of its corruption and their disdain for it, and they attempt to distance themselves from it, using instead either legitimate portuguese or english, the official language, particularly in Bombay and its suburbs.'

Interestingly enough, the prestige that SP enjoyed in Diu in light of the circumstances of colonialism - comparable to the situation just described in early 20th-century Bombay - has not been dispelled since 1961. The present vitality of this concept allows the notion of SP as an indicator of education and wealth to go on structuring the very community of native speakers of DIP. In other words, among the Diuese Christians, proficiency in SP is still an index (if not a pre-requisite) of high social status.

On the other hand, social status is clearly not the only connotation of SP. Some scholars are acutely aware of the fact that language use and social factors interact in multiple domains. This is particularly evident in the following observation:

> [n]ational, ethnic, racial, cultural, religious, age, sex, social class, educational economic, geographical, occupational and other groupings are all liable to have linguistic connotations. The degree of co-occurrence of boundaries will vary from one society to another, the perception of the degree of co-occurrence will vary from one individual to another. (LePage and Tabouret-Keller 1985:248)

An additional proposal of these authors, as evident from the previous quote, is that not only do different societies express different combinations of these domains through language use, but there is also variation concerning the interpretation of these social manifestations on an individual level. The challenge then is to produce an inventory of social variables for which allegiance to (Indo-)Portuguese is relevant and attempt to explain it with reference to past colonial structures and policies. The following have been identified through observation of distribution and use of both DIP and SP in present-day Diu:

a) Religion: this is by all accounts the social domain within which the strongest connotation of Portuguese operates, in Diu; as mentioned in section 2.1.1, Portuguese is widely considered 'the language of the Christians'[4], and there are various logical reasons for that. On the one hand, Christianity is intricately connected to colonial rule because it was propagated in the territory during that period and by its perpetrators. *Catholicism* has also been construed as a defining trait of the Portuguese, and nowadays *Portugueseness* is seen as a defining trait of the Catholics; to the extent that the Christian community has since 1961 inherited the role of upholders of this culture, so have they also inherited the association with native use of the colonial language. On the other hand, it is indeed the Catholics who speak DIP natively. Certain members of this community also claim direct descent from the European settlers.[5] Language is absolutely instrumental in setting the boundaries of the Christian community and enforcing the notion of its separate identity with regard to the rest of the population.[6] These considerations are, of course, more applicable to families which experienced Portuguese rule (be it in Diu, Daman or Goa) than to those Catholics who settled in Diu after 1961 from various parts of India (see section 2.1.1).

b) Social status: the territory's colonial history has enforced an association between the Portuguese language and administration, and more generally between the Portuguese language and a ruling elite. Administration was not exclusively carried out by Christians - at least in the period immediately preceding decolonisation - nor is it at present but, as we have seen, they remain well represented in minor administrative posts. It is perhaps not surprising that the centres of the territory's administration (Collector's office and depen-

---

[4]The identification of (restructured) Portuguese with the Christians as an ethnic division is not incidental or unique. The Portuguese-lexified Creoles of Daman, Macau and Malacca are similarly interpreted by the societies in which they are embedded; this is particularly evident from the local appellation of the creole of Malacca as well as its speakers, who are known as *Kristang* 'Christian' (see Baxter 1983, 1988).

[5]This claim is not easy to verify in all cases. The fact that the Catholics have Portuguese family names is usually invoked as evidence of descent, but it is also known that, in the past, Christian converts would normally adopt a new surname, often modelled on that of a prominent Christian figure, the missionary or someone else responsible for the conversion.

[6]It must be mentioned that, alongside religion and language, other cultural manifestations ultimately traceable to Portuguese presence are characteristic of the Diuese Christians, such as clothing, certain food items, songs and dance, etc.

dencies, court) are nowadays places where it is easy to find fluency in Portuguese, among Christians as well as non-Christians. This particular connotation of Portuguese is likely to be weakened by the fact that high administration posts now require fluency in both Hindi and English.[7] If at all, the social status association of Portuguese may be upheld through contingent notions such as wealth or education. At present, the correlation of SP with social rank also applies to the community of DIP native speakers. Although fluency in SP *per se* does not determine someone's status, it is true that the most influential among the Christian community are generally those with better knowledge of the standard norm.

c) Ideology: nostalgia for the past is not devoid of significance when addressing the maintenance of both DIP and SP. The Christians feel a particular cultural bond with Portugal, but significant nostalgia for the colonial era is also to be found among members of the Hindu and Muslim communities. This attitude towards the past is highly localised: it appears that those able to manifest themselves in the colonial language are generally those who feel particularly close to the colonial era, reflecting a particularly privileged relationship with the Portuguese social structure (among those who experienced it), either by means of employment, education or simple proximity to the foreign elite or the local Christian community. As far as the community of native speakers of DIP is concerned, it is quite clear that the younger speakers feel weaker affection for the colonial past than their elders.

d) Age: the integration of Diu in India sparked a process of cultural acculturation with very clear linguistic consequences. Reflecting the chronology of the island's decolonisation, knowledge of DIP and SP is much more widespread among the older generation, followed by the middle generation and very little among the younger generation. The only members of the younger generation with proficiency in DIP are the Christians, as this is their community's first language. Among the Muslim and Hindu children and teenagers, knowledge of either DIP or SP does not go beyond some formulas, unless they have recently experienced a period of emigration to a Portuguese-speaking region.

e) Economic affluence: the considerations made above concerning the correlation between knowledge of Portuguese and social status are intimately connected to the issue of economic affluence. Chapter 3 will make it clear that, although trading was largely dominated by local tradesmen rather than Europeans throughout the colonial period, the regulatory power of the colonial authorities meant that commercial success depended on keeping good relations with the ruling elite. On the other hand, emigration was from very early on a common route to affluence, and it still is. Crucially, many of the preferred destinations for Diuese emigrants (Mozambique, Portugal, Goa) have provided close contact with SP. The link between fluency in Portuguese and economic affluence is certainly reinforced by the fact that many of these families, whose children have in some cases grown up speaking SP, regularly return to Diu. The correlation between economic affluence and a higher degree of education is self-evident. In the case of 20th-century Diu, education

---

[7]Daman and Diu being part of a *Union Territory*, under the direct dependency of the central government, these are the official languages of bureaucracy with official institutions. Gujarati, however, does play a central role in communication with the population at large.

higher than basic had (and still has) to be attained outside the territory, and before 1961 that usually meant Goa or Portugal.

f) Education: as mentioned in section 2.1.2, Portuguese-medium (primary) education is largely responsible for the Portuguese proficiency of many members of the older generation. According to regular reports of the colonial administration, in the 19th and 20th centuries there were several Gujarati-medium schools in the territory and two Portuguese primary schools (e.g. Nunes 1898, Moura 1901), both located in Diu Town. Portuguese-medium education was not meant exclusively for the Christian population. It is unclear what the policy of admission to this school was, and therefore it is not possible to explore the interactions between Portuguese-medium education and other variables, such as economic affluence or social status. This connotation of Portuguese is less relevant for the younger generations, as Portuguese-medium education was discontinued after 1961, although Portuguese courses continue on offer until the 1980s. An anonymous report on the state of Portuguese teaching in a 1986 Goan journal[8] indicates that only one school offered Portuguese as an optional subject in Diu and that, in the academic year 1984/1985, only 5 students (all of the 9th grade) attended Portuguese classes.

2.2.1.3 Consequences of Standard Portuguese prestige

The high prestige value attributed to SP among the speakers of DIP, described in the previous section, grants it the status of a norm. What is interesting to observe is that, given the limited role of SP in present-day Diu, this norm is not entirely available to speakers of DIP. As such, it seems to be nothing but a latent force for the everyday use of DIP. Certain situations, however, may re-ignite the conscience of the prestige asymmetry, such as for instance the presence of a dignitary and/or a (native) speaker of SP.[9]

One of the concrete consequences that have been observed in such situations is heightened self-conscience and a diminution of verbal confidence, most notably on the part of those who dominate SP the least. Feelings of inferiority on the part of DIP speakers may cause them to speak extremely cautiously or ashamedly towards an interlocutor associated with SP, or not at all. Some may feel more comfortable in another language (e.g. English), so as to attenuate the social asymmetries implied by the DIP-SP contrast. It is understandably easier to gather samples of the vernacular from children or relatively unschooled speakers, as their access to SP is/was less extensive than for other members of the community; this usually correlates with a minor awareness of the differences between the two registers and less hesitation in speaking DIP. Children in particular are very unashamed about exhibiting vernacular features.[10]

---

[8] *Revista da Academia da Língua e Cultura Portuguesa*, vol. 6.12, p. 12.

[9] This fact poses important challenges to the task of language documentation, as it has the potential to enhance the effects of the inescapable *observer's paradox*. These issues are dealt with in section 4.1.

[10] This may give the impression that their register is intrinsically more 'creole-like' than that of adults. Tomás (1992:65) hypothesises that the speech of children is recreolising due by contact with Gujarati, and yet one must entertain the possibility that it is precisely the children's lesser skill or will in reverting to SP which is responsible for the perceived differences. My observation indicates that even the features which older speakers attribute exclusively to the speech of children (e.g. the semantic extension of *bunit*

In such situations, DIP speakers will revert to SP or variable approximations of it whenever possible. Some families often attempt to correct the most salient DIP features in the speech of their members, particularly th children. Many speakers attempt to implement what they consider to be corrections to their own speech. Among the most recurrent is the expression of verbal morphology to indicate person categories, which I will describe as illustration.

Observation of free-flowing DIP interaction makes it possible to firmly establish that DIP, despite its high level of diffusion, does not mark person distinctions on the verb (see section 6.2 for a full account). Obligatory expression of person through verbal morphology, on the other hand, is a very salient characteristic of SP, which may explain why this feature is so often targeted in normative self-corrections. The two contrasting sentences in (1) were uttered in quick succession by a young male speaker during an elicitation session, when asked to translate the sentence 'I don't know'. The sentence in (1a) is the standard construction in DIP, as confirmed by various other collaborators, whereas in (1b) the speaker employs a first person singular present verb form, following SP:

(1)   a.   *yo nã   sab.*
           1s NEG$_{cl}$ know.NPST
           'I don't know.'

      b.   *yo nã   sey.*
           1s NEG$_{cl}$ sey
           'I don't know.'

The crucial element here, one which confirms that person marking on the verb plays no role in DIP, is that these corrections are often implemented variably at different steps of the discourse, or they may extend beyond what would be dictated by the norm; this is the phenomenon known as *hypercorrection*. A few minutes later, the same speaker of the sentences in (1) produced the following translation of the sentence 'We don't know if Ashley is here':

(2)   *nɔs nã   sey si Ashley sta aki.*
      1p NEG$_{cl}$ sey if A.   sta here
      'We don't know if Ashley is here.'

The fact that the same form *sey* is employed in this case, in which the subject is first person plural and not singular, is an instance of hypercorrection which reveals hesitation as to the use of the targeted SP form. *Sta* is also a form modelled on SP *está*, the third person singular present form of the stage-level copula, which would normally be the DIP form *te* in unconstrained DIP interaction.[11]

'beautiful' to cover the meaning 'tasty') are to be found in both the middle and older generations as well.

[11] If I am right in interpreting this form *sta* (or *ista*) as an instance of normative correction, then it must be conceded that it is a very widespread one, as more speakers have implemented it in the course of recording sessions.

Another instance of correction in the speech of the same collaborator is given in (3). It must be the case that, for this speaker, the second person DIP possessive *duse* must be flagged as distinctly non-SP, and he therefore avoids using it when translating these two sentences. The forms he employs instead, however, are all different from each other: the two forms *tyo* and *tew* are evidently modelled on SP *teu* 'your.m.s', although in the first case SP morphology would require additional gender marking; the form *syos*, possibly from *seus* 'your.m.p' with what appears to be an SP plural suffix, would also not be allowed with a singular object; *sua* and *su* are apparently modelled on SP *sua* 'your.f.s':

(3) a. *faz   favor   chama   tyo mãy.*
         do.NPST favour call.INF *tyo* mother
         'Please call your mother.'

    b. *õd   sta tew irmãw?*
         where *sta tew* brother
         'Where is your brother?'

    c. *õd   sta syos irmãw?*
         where *sta syos* brother
         'Where is your brother?'

    d. *de   es   kart p-əkəl   piso kẽ faz   lĩ̃pez   də*
         give.IMP DEM$_p$ letter DAT-DEM$_d$ person REL make.NPST cleanliness of
         sua kaz.
         sua house
         'Give this letter to the person who cleans your house.'

    e. *yo ẽkətr-o   su mãy   oj.*
         1s meet-PST *su* mother today
         'I met your mother today.'

Various such examples could be given to illustrate the phenomenon of (hyper)correction. Recognising the phenomenon is more essential than one might imagine. Given that my purpose is to achieve an accurate grammatical description of DIP in its unconstrained type, it is imperative to be able to filter out instances of hypercorrection from the analysable corpus. On the other hand, it is essential not to mistake irregular corrections spurred by normative pressure with simple variation, which is not always an easy task. For descriptive purposes, I have either disregarded or approached very critically data from speakers who revealed a) strong hesitation and self-conscience, and/or b) widely varying or inconsistent use of certain constructions/forms (in particular if the alternatives involved showed clear influence of SP). See chapter 4 for further details.

2.2.2 Diu Indo-Portuguese and Daman Indo-Portuguese

The connection between Diu and Daman is a long-standing one, not only on account of their relative physical proximity but also because, ever since they both came under Portuguese control, they have always been part of the same political unit. In the early days of the *Estado da Índia*, the administrative unit known as the *Província do Norte* 'Northern Province' (see section 3.1) - which included a stretch of coast from Daman to Chaul - extended to include Diu across the Gulf of Cambay. After the integration into India in 1961, the three former Portuguese colonies of Goa, Daman, Diu, Dadra and Nagar-Haveli integrated a single Union Territory governed from Goa. With the proclamation of the state of Goa in 1987, the Union Territory was reduced to Daman, Diu, Dadra and Nagar-Haveli.

The crucial issue for us is that this political closeness has been accompanied by strong cultural and linguistic ties. Virtually every Christian family in Diu extends to Daman, as a result of the old practice of intermarriage between the Catholics of both territories. At present, several native Diuese live in Daman and vice-versa, and it is common for people from one territory to visit friends and relatives in the other. One should not think, however, that population interchange is balanced. Daman, with a vibrant industrial scene, is much more attractive to the Diuese than Diu is to the Damanese. Furthermore, the Catholic community of Daman (and therefore the community of native speakers of Daman Indo-Portuguese) is much larger than that of Diu, numbering around 4000 (Clements and Koontz-Garboden 2002). Most of the Damanese residing in Diu, whether or not married to local Diuese, have been assigned there on some official duty.

Linguistically speaking, the two territories also share a number of similarities. To begin with, they are both located within the Gujarati-speaking area, surrounded by the state of Gujarat. On the other hand, the Indo-Portuguese varieties of Diu and Daman are mutually intelligible and very similar in many respects. The strong similarity between the two varieties is commonplace among their speakers, not just at present but in the past as well. In his 1903 study of Daman IP, Dalgado (an external observer) makes the following observation:

> O dialecto de Damão, junto com o de Diu, *de que muito se aproxima*, pode bem ser considerado como um dos subdialectos do crioulo *norteiro* ou, como é denominado na Índia, português dos norteiros, que os tem muitos, com variantes de maior ou menor importância, na presidência de Bombaim. (my emphasis, Dalgado 1903)

> 'The Daman dialect, together with the one from Diu, *to which it is very similar*, can properly be considered one of the subdialects of the *norteiro* creole or, as it is known in India, portuguese of the norteiros, of which there are many, with more or less important variants, in the Bombay presidency.'

As a result, and considering that DIP is rather a diffuse variety with little normative power, one begins to understand why speakers of Daman Indo-Portuguese residing in Diu do not necessarily need to adjust to DIP, even if they eventually do.[12]

---

[12] Unfortunately, I am not acquainted with the situation of the Diuese residing in Daman and whether

The apparent lack of normative pressure imposed on Daman IP speakers by DIP speakers in Diu gives rise to interesting patterns of variation which reveal strong family effects in language transmission. One consistent grammatical difference between DIP and Daman IP, one which is even pointed out by DIP speakers as a stereotypical trait of Daman IP, refers in particular to the expression of progressive actions: DIP employs an infinitival verb form modified by an imperfective auxiliary (4a), whereas Daman IP makes use of a gerundive verb form also modified by a similar auxiliary (4b). Example (4b) was taken from Clements and Koontz-Garboden (2002:215), including original orthography and glosses ('PRES' stands for Present and 'PROG' for Progressive):

(4)   a.   *Leslie tə   kãt-a.*
             Leslie IPFV.NPST sing-INF
             'Leslie is singing.'

        b.   *Joyce te   kantan   agɔr.*
             Joyce PRES sing-PROG now
             'Joyce is singing right now.'

Gerund forms - some of which do occur marginally in DIP (see section 7.1.5) - are constructed with the suffix *-n*. The corpus recorded in Diu for the present description did contain several instances of gerundive forms expressing progressive actions, some of which are transcribed below:

(5)   a.   *Armando kai-w   i   pə   el te   chur-an.*
             A.   fall-PST and DAT 3s IPFV.NPST cry-PROG
             'Armando fell and he is crying.'

        b.   *i   yo te   kure-n.*
             and 1s IPFV.NPST run-PROG
             'And I am running.'

It turns out that the speakers who employ gerundive verb forms are either Daman natives or direct descendants of Damanese. This observation simultaneously reveals two interesting facts concerning the relationship of DIP and Daman IP, viz. that the two codes can coexist relatively stably within the same territory and small social circle, and also that families play a decisive role in transmitting language, even when their variety is not dominant in their environment.[13]

One further issue surrounding the relationship between the two varieties of IP has to do with the very frequent claim, pronounced as a truism and with much pride by speakers

---

or not they feel much pressure to assimilate Daman Indo-Portuguese features into their speech, so that it is not possible to provide the logical counterpoint here.

[13]See also section 7.9.1 and chapter 7 fn. 37 for a similar case involving a particular discourse marker.

of DIP, that their variety is closer to SP than Daman IP is. It is difficult to measure the truth of this perception, but it is interesting that we can find close repetitions of the claim in the literature:

> Aí, tal como em Goa (e muito mais do que em Diu), verifiquei que a situação linguística actual resultou, antes do mais, dos «processos de confronto» e de «aculturação (recíproca)» entre a língua portuguesa, nos seus variados dialectos naquelas paragens, e outras línguas que prolifera(va)m na costa hindustânica, tocada muito particularmente pelas nossas gentes e pela nossa cultura. (Matos 1987:327)

> 'There [Daman], just as in Goa (and much more than in Diu), I have verified that the present linguistic situation resulted, above all, from the «processes of confrontation» and «(reciprocal) acculturation» between the portuguese language, in its various dialects of those lands, and other languages which proliferate(d) on the hindustani coast, especially touched by our people and our culture.'

One possible reason why a visitor, upon superficial observation, might have this impression may have to do with the vast quantitative difference between the speakers of DIP and those of Daman IP. Another possibility might be that, overall, the native speaker community of DIP might be more proficient in SP than the speakers of Daman IP, which says nothing about the typological distance between SP and the language they speak in unconstrained situations. Such claims raise more doubt than certainty, and accurate comparative research is required before this myth can be either debunked or confirmed.

2.2.3   Diu Indo-Portuguese and the languages of India

Of all the languages officially recognised in India, three are particularly relevant in the context of modern Diu: a) Gujarati, the state language of the surrounding state of Gujarat, b) Hindi, the national language of India, and c) English, recognised as an official language of the country.

Although Diu integrates a separate administrative unit, the island is part of the cultural landscape of the state of Gujarat. Gujarati is the native language of the Hindu and Muslim communities, and was so during Portuguese rule as well. Classes in the government-run schools in the territory are taught in Gujarati and in English-medium schools Gujarati is still one of the compulsory courses. The DIP-speaking population is extremely fluent in Gujarati, as it is necessary for daily life, social interaction and commercial exchange. The fact that Gujarati is an additional common language to all native speakers of DIP (from early childhood) results in a particularly high resort to the Gujarati lexicon in DIP discourse. The first illustrative example below was uttered by a male child, who employs the Gujarati word *kanjus* 'avaricious, miser' within a DIP matrix; (6b) includes the Gujarati grammatical word *avŭ* 'so, thus'; (6c), spoken by a female child, exemplifies the participation of the Gujarati word *pechŭ* 'squashy' in a DIP conjunct verb (see section 7.1.4 for a definition of the term and full description):

(6) a. *el fal-o   yo ɛ*   kanjus.
            3s say-PST 1s COP$_i$.NPST *kanjus*
            'He said: "I am avaricious".'

      b. *es   ɛ*   avũ *piken.*
            DEM$_p$ COP$_i$.NPST *avũ* small
            'It is this small'.

      c. *el tə   faze* pech.
            3s IPFV.NPST make *pech*
            'He is squashing (it).'

Given that Gujarati was an essential participant in the contact situation resulting in DIP, and considering that contact with Gujarati has been a historical constant, it is not surprising to encounter striking parallels between the two languages on various levels (phonological, syntactic, lexical, pragmatic). These are dealt with in the chapter 9, but with regard to the lexicon it is important to concede at this point that it is often difficult to ascertain whether or not a word of ultimate Gujarati origin - such as the ones in (6) - should be treated as a borrowing or as a component of the DIP lexicon. Seeing as we are dealing with the linguistic production of members of a speech community which is at least bilingual in DIP and Gujarati, this is perhaps a non-question, even normative in nature. Yet, in practical terms, it is inescapable if our purpose is that of describing one of the two codes proficiently used by the community. Chapter 8 indicates some methodological considerations adhered to in response to this issue and approaches the issue from an ecological perspective according to which lexemes from various sources, often semantically equivalent, constitute a relatively unstructured pool from which the speakers are able to select during linguistic production.

Hindi, though not indigenous to the area, has a very prominent role in the Indian media and the national discourse. Although virtually all the speakers of DIP have at least passive knowledge of Hindi, it is the young and middle generations in particular who tend to master this language, also taught as a compulsory subject in school. A similar situation obtains concerning English, which has firmly taken up the role of a pan-Indian language; in Diu it attracts strong allegiance among the DIP-speaking children and youngsters (its role in nationwide business, as well as tourism locally, is perhaps involved in this preference). All the children, teenagers and the middle generation who speak DIP natively attend or have attended the English-medium Catholic school, and they are increasingly exposed to this language in religious ceremonies as well. As such, these speakers attain extremely high proficiency in (Indian) English. Just like Gujarati, and much more than Hindi, English provides many lexical items to the everyday discourse of DIP speakers. Consider for instance the following utterance, in which a young female collaborator circa 30 attempts to retrieve a Portuguese word and, failing to do so, resorts to an English equivalent:

(7) *ũ di faz-e   es* bridge, *uki dig?*
     one day make-INF DEM *bridge* what say.NPST

'One day they made this *bridge*, how do you say it?'

The particular adhesion of DIP speakers to English is perhaps not unexpected, even if one discounts the obvious economic advantages it promises in modern-day India. All over the country, Christian communities traditionally show the same tendency to use English not merely as a professional tool but also as an in-group language, which grants them particularly high proficiency in this language.[14] This is true of the Anglo-Indian communities, most prominent in South India, some of which were originally Indo-Portuguese speaking (e.g. in Cannanore and Cochin). More recently, the same can be observed among the Catholics of Goa, and it is particularly telling that in both Daman and Diu church services are now mostly conducted in English rather than Portuguese or Gujarati. A related observation is that, although English-medium schools and colleges now proliferate in India, the earliest and most prestigious such institutions are Christian-run. This is the case, in Diu, of the Nirmala Mata Catholic school, which is attended by all DIP-speaking children. In the case of modern-day Diu, it will not be surprising if, as SP and DIP partially lose vitality, English should come to replace them in various domains, as this language retains the same associations of foreignness and, perhaps, Europeanness.

2.3 Variation

The distinction between *variation* and *(hyper)correction* is an essential one. By (hyper)correction, I refer to the type of changes implemented by speakers in their own speech under normative pressure, typically in interaction with speakers of SP, often resulting in inconsistent and hesitant production of the type described in more detail in 2.2.1.3. I interpret variation, on the other hand, as the spontaneous and natural selection of competing linguistic forms, structures and strategies, either by different speakers or in the speech of a single speaker. This does not mean that normative pressure is irrelevant to issues of variation. In fact, out of the pool of (nearly) equivalent possibilities available to a given speaker, some may have more or less strong associations with a norm and therefore see their selectability enhanced whenever external circumstances dictate that normative pressure applies.

Given the heterogeneous multilingual environment in which DIP evolved and presently operates, it is not surprising that individual speakers have at their disposal a highly complex pool of competing possibilities from different sources. Furthermore, the fact that SP congregates the characteristics of linguistic norm - SP being something of a removed (and unknown, for many) standard - in turn means there is no local norm, and the result is a rather diffuse linguistic variety.

Even if one discounts the fact that the vast majority of the Diuese population has always been Gujarati-speaking, normal linguistic interaction on the island must always

---

[14]This preference for English is not a prerogative of the Christian communities in India, but can be observed among other minorities who need to highlight separate identity or consider themselves in some fundamental way displaced. The case of the Parsi community is a classical example, as is that of the Marwari diaspora.

have included several varieties of Portuguese (including SP and other varieties of Indo-Portuguese), several degrees of proficiency in the language, L2 types of transfer, formulaic speech, as well as the different registers available to the native speakers themselves. Variation, therefore, is likely to have been the norm throughout the history of the language.

The descriptive chapters discuss various instances of modern-day variation, but I will mention some here as illustration of what has just been said. The first example concerns NP-coordination (see 7.9.1), for which DIP can employ any of the following conjunctions: a) *i* (from Ptg. *e* 'and'), also used for clausal coordination (8a); b) *ku* (from Ptg. *com* 'with'), also a comitative and instrumental marker, disallowed in clausal coordination(8b); and, rarely, c) *may* (from Ptg. *mais* 'together with, (more)'), similar to the disjunctive co-ordinator and a comparative quantifier/intensifier, disallowed in clausal coordination (8c):

(8) a. *Conchita i Balfina ε   mεdroz.*
         C.   *i* B.   COP$_I$.NPST fearful
         'Conchita and Balfina are very easily frightened.'

     b. *crocodile ku makak, Steven kõt-o   istɔr.*
         crocodile *ku* monkey S.   tell-PST story
         'Steven told the story of the crocodile and the monkey.'

     c. *yo may doy rapas tə   trabəy-a.*
         1s *may* two boy   IPFV.NPST work-INF
         'Me and two boys work (here).'

The option for one or the other does not depend on the syntactic role of the entities making up the NP. To illustrate the co-availability of the various forms to the same speaker, a short dialogue is transcribed in (9) showcasing the free distribution of *i* and *ku*. *C* signals the utterances of the collaborator, in this case a male teenager, and *R* refers to the researcher, who asks for clarification of a particular sentence:

(9) *C* nə   *Go* yo te   *bastãt* cousin *i* auntie.
         LOC Goa 1s have.NPST many cousin *i* auntie
         'I have many cousins and aunties in Goa.'

     R [te bastãt . . .?]

     C *cousin ku auntie.*
         cousin *ku* auntie
         'Cousins and aunties.'

In the case of NP coordination, it seems that speakers are familiar with all possible strategies (though perhaps with idiolectal preferences), so that communicative success is

not hindered by any particular choice. The same holds true for the variation observed in the form of indefinite pro-forms (see 6.6.5). DIP has two distinct series of indefinite pro-forms, one negative and one non-negative. As far as non-negative indefinite pro-forms are concerned, the only self-standing pro-form is the pronoun *aŋe* 'somebody', whereas all other instances of indefinite reference employ an analytical structure consisting of an indefinite element and a generic nominal (such as 'time', 'place', 'people' or 'thing'). The indefinite element may be one of three: a) *aŋu*, from Ptg. *algum* 'some', by the far the most common of the three possibilities; b) *kwɔlki*, from Ptg. *qualquer* 'any'; or c) *sɛrt*, from Ptg. *certo* 'a certain', the least common of the three. For the purpose of forming indefinite pro-forms, the three are equivalent, and the selection of one or the other is purely arbitrary. The elicitation of the same sentence from various collaborators resulted in the variation recorded in (10):

(10) *aŋu/kwɔlki/sɛrt jẽt   pəd   abr-i   kok   sẽ   fak.*
      some   people can.NPST open-INF coconut without knife.
      'Some people can open coconuts without a knife.'

One final example of variation, on a phonetic level, concerns the optional insertion of a glide preceding stressed mid-high and mid-low vowels: a palatal approximant [j] before front vowels [ɛ] and [e], and a labial-velar approximant [w] before back vowels [ɔ] and [o] (see 5.1.2). The words below (given in their phonological form) therefore admit two phonetic realisations:

- /papɛl/ 'paper' -   [pa'pɛl]
                      [pa'pjɛl]

- /mem/ 'EMPH' -   [mem]
                      [mjem]

- /fɔrt/ 'strong' -   [fɔrt̚]
                      [fwɔrt̚]

- /kojz/ 'thing' -   [kojs]
                      [kwojs]

These are not instances of geographical variation; the physical territory of DIP is much too small to allow for regional variation, in particular since the Christian population concentrates exclusively in Diu Town.[15] The examples provided in this section show that

---

[15] On the other hand, if one were to consider DIP and Daman IP to be variants of the same language, it would result that the language does indeed show regional variation. In addition, that variation would manifest itself in the very territory of DIP through the influx of speakers of Daman IP; cf. also the considerations around the instances of morphological progressive marking in the corpus of spoken language collected in Diu, section 2.2.2.

variation can be widespread even in a small, tight-knit language community, affecting such seemingly stable elements as grammatical items.

2.4 Vitality and maintenance

The observed vitality of DIP and the size of the L1-speaking community encountered is in contrast with previous accounts, which classified the language as extinct (v. Smith 1995) or spoken only by a very small number of people (Tomás 1992), estimated at around 15 in Clements (1991:351). Nonetheless, the future of the language is still quite unpredictable. In this section, I will explore some of the factors which may impact positively or negatively on the language's long-term maintenance, based on my observation of patterns of language use, as well as sociolinguistic and official attitudes to the language.[16]

One preliminary observation is that the long-term survival of DIP among its L1 speakers will probably profit from the maintenance not just of cultural ties with Portugal and the Portuguese-speaking world, but also of a more or less prominent role of SP in the territory. The Portuguese language and culture, whether encountered in religious ceremonies, through contact with foreign and emigrant visitors or otherwise, is an essential component of the Catholics' sense of identity, as is their particular connection to both Daman and Goa. On the other hand, L1 speakers of DIP declare that, if they were to have access to education in SP, they would switch and adopt the norm in detriment of DIP. One should not be surprised at this claim, in view of the prestige asymmetry between DIP and SP described in 2.2.1.3, but it is difficult to predict what actual threat such a situation would pose to the survival of DIP. The local code is, after all, endowed with considerable affective value, which is favourable to its maintenance among the Catholics.

Another favourable factor is that, as a result of the above, the transmission of DIP has not been discontinued. Most Catholic families still prefer DIP as their in-house and in-group language. As such, the children acquire DIP as their native language, even before they begin to acquire Gujarati (which happens very early on) or English (upon starting their education).

Considering the dearth of native speakers, the closeness of the ties linking Diu and Daman, in which Daman IP is spoken by nearly 4000 people, is crucial. The role of Goa in this respect is less significant for two reasons. On the one hand, among the Goan Catholics, Konkani is much more prominent as a native language than Portuguese is. In addition, the population exchange between Goa and Diu is presently less significant than that between Diu and Daman.

The reduced size of the native speaker population is perhaps the greatest challenge to the survival of DIP: at less than 200, the community is too small to ensure stability; it would be vulnerable, for instance, to a significant wave of migration. In addition, despite a moderate influx of people from other regions, the Christian population in Diu has been progressively declining for over a century. With respect to this issue, it is interesting to notice the official report of the Governor of Diu, João Herculano Rodrigues de Moura, concerning the year 1900. The statistics provided in this document clarify the situation

---

[16] For suggestions of concrete interventions aimed at potentiating maintenance, see Cardoso (2006a).

concerning the Christians:

> O total da população catholica do districto é de 361 almas, 30% das quaes não são naturaes de Diu, são militares e funcionários de Goa aqui destacados. [. . .] devo considerar 240 christãos naturaes de Diu para a sua população catholica.

> 'The total catholic population of the district numbers 361 souls, 30% of which are not natives of Diu, they are military and administrative personnel from Goa assigned to come here. [. . .] among the catholic population I should count 240 christians as natives of Diu.' (Moura 1901:40)

Considering the present Catholic population of Diu, numbering *circa* 250 out of which less than 200 are natives of the territory, the decrease of the community looks rather slight. But the document makes another interesting comment which reveals the trend had started earlier than 1900 (see section 3.2.3.1):

> A marcha da população christã é em Diu negativa, apezar de se considerar os obitos de cholera á parte. A raça não está parada, tende a desapparecer, e em um futuro não muito longo. Continuando as cousas como até aqui em menos de 20 annos o *norteiro* de Diu será para a ethnographia o que é hoje o *máoris* da Nova Zelandia!

> 'The growth of the christian population in Diu is negative, even if we count the cholera casualties separately. The race is not static, it tends to disappear, and not in such a faraway future. If things remain as they are, within 20 years the Diuese *norteiro* will become, for ethnography, what the New Zealand *maoris* are nowadays!' (Moura 1901:40-42)

The survival of DIP among non-Christians is an entirely different matter. With no effective transmission onto the youngsters, knowledge of DIP seems destined to die out within two generations at most. Some DIP or SP formulae may last longer, but no fluency is to be found among the younger Muslims and Hindus. The influx of SP through migration to and from Portugal and Mozambique (in particular among the Hindu population) may secure some vitality of this variety in the territory for a long time, but DIP is likely to be abandoned.

In Diu, the institutional relevance of both DIP and SP has been minimal since decolonisation.[17] The colonial undertones of the Portuguese language, which affect DIP by

---

[17] Speakers and policy-makers are allowed some degree of manipulation over a language's ecology, which is particularly evident in moments of power reversal such as decolonisation; whether the process is peaceful or not makes no definitive predictions as to the degree of support the new institutions will assign the former colonial language. One of the most decisive variables for the maintenance of an endangered language is indeed official recognition. Giles et al. (1977) include *institutional support* in the inventory of factors contributing to language vitality, alongside demographic and other sociolinguistic considerations:

> Institutional Support variables refer to the extent to which a language group receives formal and informal representation in the various institutions of a nation, region or community. The vitality of a linguistic minority seems to be related to the degree its language is used in various institutions of the government, church, business and so forth. (Giles et al. 1977:309)

proxy, have resulted in a very low level of interest - not to say an active effort of demotion - on the part of the political authorities. With no official role in the administration of Diu, Portuguese presently plays a role in one institution only, viz. the Catholic Church. Most services and special celebrations were traditionally conducted in Portuguese (SP for the most part); recently, however, the parish has seen a period of services exclusively in English. The main reasons behind this change had to do with the appointment of a new generation of Goan-educated priests with no knowledge of Portuguese, as well as the need to accommodate the non-Portuguese-speaking section of the Catholic community. Portuguese liturgy was recently resumed, by popular demand, alongside more frequent services in English.

There seems to be a renewed interest in minority languages in India (see Annamalai 2003, Abbi 2008, Bhatt and Mahboob 2008), which is helping fine-tune the country's language inventory. Romaine reports a revealing discrepancy in the most recent national censuses:

> In the 1981 census in India 107 mother tongues were reported. Only 20 years later, however, 1,652 mother tongues were reported. The discrepancies here are due to a number of factors. One is that a given mother tongue may be called by as many as 47 different names depending on the ethnic, religious and other affiliation of the person who claims it. Out of all these varieties, however, only 15 mother tongues are recognized as 'major languages' by the Indian government. (Romaine 1995a:27)

Given that minority languages are particularly vulnerable as far as their long-term survival is concerned, this new trend is, from a pluralistic perspective, undoubtedly positive. With respect to its consequences for the official recognition and maintenance of DIP, one needs to be less optimistic. In order for DIP to profit from this new interest, it will need to be dissociated in the official and scientific discourse from Portuguese, with its negative colonial echoes, and legitimised as a uniquely Indian language. This constitutes an additional hurdle on the way to acceptance which most minority languages of the country do not face.

It is clear that the long-term maintenance of DIP is faced with significant frailties. If we take into account the various indicators discussed in this section, it emerges that DIP, notwithstanding present signs of vitality and continued transmission, must be treated as a seriously endangered language.

Chapter 3

Social history

As a preamble to the present section, the timeline below indicates some key dates and events in the history of Diu.

Late 15th-century	Diu integrates the Sultanate of Gujarat.
1509	Battle of Diu.
1513	A short-lived Portuguese trading post is founded.
1535	The Portuguese are allowed to build a fortress in Diu.
1537	The city and port are briefly under Portuguese control.
1538 (June-November)	First siege of Diu.
1546 (March-November)	Second siege of Diu.
1554	The Portuguese take complete control over the city and the entire island.
1612	Commercial competition from Surat; beginning of the commercial decline of Diu.
1654	Dutch armada enters the port of Diu.
1668	Plundering by Omani forces.
1715	Annexation of Simbor.
1718	The walls of Diu fort are rebuilt.
1796	French bombardment of the island.
1961	Military take-over of the island by Indian Union forces.

Table 3.1: Key events in the colonial history of Diu

This short summary is a preliminary overview of the complex (colonial) history of the island, which is explored in detail in section 3.1. After describing the main events in Diuese history, from its rise as a major trading port to the decolonisation in 1961, section 3.2.2 will characterise the social relations which obtained in the *Estado da Índia* in general - and Diu in particular. Based on the available evidence, section 3.3 will then attempt to reconstruct the sociolinguistic conditions which resulted in the formation of Diu Indo-Portuguese, its development and present-day distribution.

3.1 History of the settlement

The short history of the settlement given here covers a rather vast period, from the time prior to Portuguese occupation to the integration in modern-day India in 1961. For convenience, this section is divided into subsections dealing with the various phases of the island's history: the early history of Diu up until the arrival of the Portuguese (section 3.1.1), the extended struggle for power from the beginning of the 16th century (in 3.1.2), the early days of Portuguese control (section 3.1.3), the subsequent decline of the island and the city on the regional stage, from the 17th century onwards (in section 3.1.4) and, finally, the process of decolonisation (section 3.1.5).

3.1.1 Diu before the Portuguese

In pre-Portuguese times, the island of Diu was known as *Diva*, derived from the Sanskrit name *Dvipa*, meaning 'island' (Altekar 1924-1925). The development of Diu as a major trading port was spurred by dynamic trade operations in the Persian Gulf area from the 9th century AD onwards, and in particular by the rise of nearby Cambay (modern-day Khambhat) in the same period.
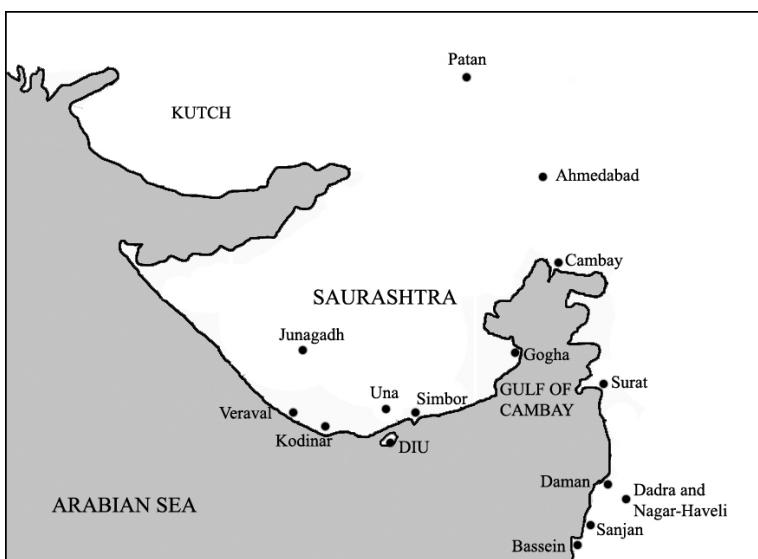
Map 3.1: Map of Gujarat and the Gulf of Cambay

There is evidence of trading routes linking Diu not only to the Arabian Peninsula but also to Southeast Asia (Vardarjan 1989). Cambay remained the dominant port of the entire region and the hub of Gujarati commercial activity for many centuries, well into the European colonial period until it was supplanted by Surat. That can be inferred from the

description made by Pyrard de Laval, a French traveler writing in the early 17th century:

> La ville de Cambay est l'une des grands et des plus riches de la côte des Indes, où abordent des marchands de tous les quartiers du monde. La langue de tout ces pays-là, comme aussi de tous les autres du grand Moghol, de Bengale et des circonvoisins, est [la] langue de Gujarat, qui est la plus grande, utile, étendue et qui s'étend en plus de divers endroits qu'aucune autre des Indes. (Laval 1998:752)

> 'The city of Cambay is one of the biggest and richest of the Indian coast, where merchants from all the corners of the world make port. The language of all that region, as well as of all the others of the great Moghul, Bengal and neighbouring areas, is the language of Gujarat, which is the largest, most useful, most extended and spread in more different places than any other in the Indies.'

Navigation in the Gulf of Cambay was however complicated by the extreme tidal bore and progressive silting, which prevented big ships from safely entering. This favoured the emergence of Diu as a trading port, as it became common for goods to be transferred there (and, later on, also in Gogha, see Map 3.1) into smaller vessels which would be able to carry them to Cambay. In time, more and more of the trade began to be carried out in Diu itself to avoid the trouble of transporting the goods to Cambay, which meant that, around 1500, the island was already an extremely important trading post (Tibbetts 1971).[1]

Commerce from Aden to Gujarati ports was partly dominated by Muslim merchants but, as pointed out by Pearson (1976:10) concerning the historical views on trade in Asia, a largely 'Euro-centric stress on trade to Europe ignores the vast bulk of Asian trade, that which was *not* bound for Europe. And in this inter-Asian trade, it was Gujaratis, not Arabs, who were dominant'. In a sense, Gujarati ports specialised in linking the commercial centres of Aden, to the west, and Malacca, to the east. Furthermore, the area was the single most important centre of cotton cloth production at the time.

The commercial prominence of Diu did not suffer from the frequent political reconfigurations of the region. Proof of this is that, when the newly-formed Sultanate of Gujarat annexed Junagadh (to which territory Diu belonged), towards the end of the 15th century, the seat of the regional administration was installed in Diu rather then Junagadh. The first appointed governor was Malik Ayaz, who 'did a great deal to nurture Diu into a premier station of trade' (Vardarjan 1989:369) and from whom the Portuguese repeatedly attempted to wrest the territory.

Diu was an extremely cosmopolitan port, and the *Tombo de Diu* (Pais 1592:fl 42) records that before the Portuguese period the city counted 'muitos rumes, abexins e fartaquyns' [many *rumes*, *abexins* and *fartaquyns*].[2] Apart from the local Indian communities (of various religious backgrounds), it is then clear that, prior to the arrival of the Portuguese, Diu hosted a significant population from the Middle East generally identified as *Rumes*

---

[1] Diu is mentioned in old sources as one of the three ports (the other being Fakaner in Southern India and Sumatra) where traders were compelled to enter and pay duties, before the Portuguese installed this system (v. Pearson 1976:15-16).

[2] For a definition of the appellations *rumes* and *abexins*, see below (this section). According to Matos (1999:54), *fartaquyns* referred to the inhabitants of the Ra's al-Fartaq Cape region, i.e. the southeastern shore of Arabia, comprising parts of modern-day Yemen and Oman.

or *Rumji*. There is some debate concerning the correct interpretation of these terms in 16th-century Portuguese sources, even though the authors themselves often attempted an interpretation of the term. Orta (1563), for instance, admits to never having been given a conclusive explanation of the distinction between a 'Rume' and a 'Turk', while Couto (1602) states that the former differed from the latter in that they originated from Thracia and Romania. For all this hesitation, the term was, at the time, generally interpreted to refer to an inhabitant of the Ottoman empire (Özbaran 2001), of whom many had commercial and strategic interests in Western India. The area of Diu in particular is very often associated to this community; according to the account of an Italian visitor in 1504 (quoted in Singh et al. 1994:5), Diu was at the time referred to as *Bundar Atturk* 'Port of the Turks', while 'it is described as "Bandar Afarmi" in the writings of Muslim scholars', in a clear allusion to Turkish presence. In 16th-century Portuguese sources, the village of Goghla (see Map 1.2) was known to the Portuguese as *Villa dos Rumes* 'village of the Rumes', and in 1510 the viceroy D. Afonso de Albuquerque informs the king of Portugal of his intention to prepare the fleet to engage the Rumes who were at Diu (see Albuquerque 1774). Given that the Ottoman empire had interests in Diu, it is not surprising that, after the Portuguese achieved a foothold on the island, the Turkish and Egyptian rulers moved a strong military offensive against them - for which, see section 3.1.2.

Another community with a long-standing link to Diu is that presently known as *Siddhis* but, throughout its history, also as *Habshis* (*Abexins* and *Abexiis* in Portuguese sources) or *Kaffirs*. These were Africans who mostly arrived in India as slaves, through channels which predate the arrival of European influence to the region. Pankhurst (2003) provides evidence that slaves were being carried from the territory of modern-day Ethiopia into Western India as early as the 1st century AD. The documentary evidence is much more robust after the 13th-century; it becomes clear that, at the time, slaves were imported from Abyssinia (from whose Arabic name, *Habash*, the term *Habshi* appears to have derived) via Arabia, most eventually converting to Islam and ending up particularly in Gujarat and the Gulf of Cambay or further south on the Malabar coast and Ceylon. The history of the Siddhi communities of India is rather complex (see e.g. Chauhan 1995, Pinto 1992, Jayasuriya and Pankhurst 2003), but it is important to recognise their significant presence in Diu prior to the Portuguese occupation, during Portuguese rule and also at present. It is certainly interesting that, as documented in early Portuguese maps of Diu (e.g., Map 3.2, by the cartographer João Teixeira Albernaz I), one of the three doors in the city walls was known to the Portuguese as *Porta dos Abaxis* 'Door of the Habshis'.

At present, the Diuese (and Gujarati) Siddhis are Muslim (see Lodhi 1992) and retain a strong cultural and communal identity. It is crucial to distinguish between this community from the Africans transported to India by the Portuguese as slaves. For this issue, see section 3.2.1.

Diu was heavily defended even before the arrival of the Portuguese, as noticed by various coeval observers. It is reported that a Solanki[3] ruler built a fortress in Diu at the turn of the 12th-/13th-century (Vardarjan 1989:369). The sea fortress protecting the entrance of the port, known to the Portuguese as *Fortim do Mar* (labelled *F. do Mar* in Map 3.3),

---

[3]A Hindu clan which ruled much of Western India between the 10th and the 13th centuries AD. The area of Gujarat, which they ruled from their seat in Patan (north of modern-day Ahmedabad, see Map 3.1), was under their control until the late 13th-century.

Figure 3.2: Map of Diu by João Teixeira Albernaz I (c. 1648)

also pre-dates European rule (Dias 2002:6).

3.1.2   The struggle for Diu

As soon as the Portuguese navigator Vasco da Gama reached the South Indian port of Calicut in 1498, thereby opening up a maritime trading route around the Cape of Good Hope with the potential to rival the traditional land routes into Europe, attempts at establishing a firm Portuguese presence in India began immediately. Although those efforts initially centred on the southwest coast of the subcontinent, inflamed reports of the strength and riches of Diu reveal that the city was coveted by the newcomers from very early on. During its first impulse, the Portuguese empire in the East did not so much aim at the acquisition of land as at the control of vantage points from which to enforce their 'sovereignty of the sea' (Newitt 2005:74-75). Their domination in the area should be based on maritime military power (therefore a permanent Armada had to be established very early on in India) allowing Portugal the right to issue *cartazes*[4] to all vessels sailing on Indian waters.

---

[4]Obligatory safe-conducts which all vessels should obtain from the Portuguese authorities in order to safely navigate the waters under their control.

Given the vitality of the Gujarati trade, it became clear that '[s]ome fortified position in the Gujerat area was certainly needed if the Portuguese system of issuing cartazes and levying customs dues was to function at all effectively, for this was one of the busiest commercial regions of the Indian Ocean' (Newitt 2005:116). It is clear that Diu drew Portugal's attention very early on when the governor of Diu, Malik Ayaz, attacked the Portuguese fleet in Chaul in 1508, during which battle the son of the viceroy D. Francisco de Almeida was killed (Morais 1997, Leão 1996). As a retaliation, the so-called *Battle of Diu* ensued in 1509, when the Portuguese engaged the combined fleets of Malik Ayaz, Egypt[5] and Calicut off the coast of Diu and vanquished them - with this victory, Portugal achieved notoriety in the region and was able to negotiate a peace treaty under humiliating circumstances for Malik Ayaz, or Melik-Ias after some Portuguese sources (cf. Leão 1996:103, 104). In 1513, the viceroy Afonso de Albuquerque was authorised to open a trading post on the island, although by 1528 the two kingdoms were again at war.

The most outstanding proof of Portugal's interest in dominating Diu is the fact that, as the viceroy D. Nuno da Cunha arrived in India to begin his mandate in 1529, his primary task - as commanded by the king D. João III - was to build a fortress on the island. Concerning Nuno da Cunha, the chronicler João de Barros says 'ao qual el Rei não mandava à India à outra cousa senão à tomar Dio' [whom the King had sent to India for nothing else but to conquer Diu] (Barros 1615:book 2, ch XV). A consistent campaign to conquer Diu with the personal involvement of the viceroy himself was launched by D. Nuno da Cunha from the 26th of February 1531 onwards. Several setbacks ensued because, as the Portuguese aggression stepped up, so did the aid sent by the Turks and Egyptians to the Sultan of Cambay in order to safeguard their trade.

The Sultan of Gujarat, Bahadur Shah, attempted to buy peace and the protection of the Portuguese by offering the territories of Bassein (see Map 3.1) and the area of modern-day Bombay (incl. Thane, Salsette, Bandra, Elephanta, Mahim and Bombay) in 1534 (Biker 1881, Leão 1996), but he was not willing to give Diu away. The situation changed in 1535 when, threatened by the Moghul expansion from Delhi, Bahadur authorised the Portuguese to build a fort in Diu in return for military aid. Below is a revealing extract of the contract (transcribed in Biker 1881):

> aquillo que vos era necesario, que em tantos anos nunqua se pôde cumprir, nem vos ouvera de vir a mão tão asynha, hum lugar pera estarem os portugueses em dio, da banda donde quiserdes, vós o mandastes pedir; eu vos ffaço mercê d'elle com estas condições
>
> [...]
>
> Item o soltão badur he contente de dar a elRey de portuguall hũa fortaleza em dio, em qualquer luguar que o guovernador nuno da cunha quiser, da banda dos baluartes do maar e da

terra, da grandura que lhe bem parecer; e asy o baluarte do maar.

[...]

E com condição que elRey de portuguall nom teraa em dio nhuns direitos nem rendas, que so a dita fortaleza e baluartes; e todos os direitos, rendas e jurdição da gente da terra sera do dito soltão badur. [...] E com condição que querendo se fazer alguns mouros da terra do soltão badur cristãos, que o gouernador o não consinta; e asy elle não consentiraa fazerse nhum cristão, mouro.

'that which you required, which did not happen in so many years, and would not be granted to you so easily, a place for the portuguese to be in diu, whatever part you choose, you have demanded; I grant it to you under these conditions [...] Item. sultan bahadur is willing to give the king of portugal a fortress in diu, any part that the governor nuno da cunha decides, towards the strongholds of sea and land, as big as you find fit; and also the sea fortress. [...] And with the condition that the king of portugal shall not have the right to any profits and taxes, but only the aforementioned fortress and strongholds; and all the rights, tax and jurisdiction of the people of the land remain with the aforementioned sultan bahadur [...] And with the condition that if any moors of the land of sultan bahadur wish to become christian, the governor shall not allow it; and likewise he shall not allow any christian to become a moor.'

Although the concession was highly constrained and anything but profitable for the Portuguese, the achievement was so crucial that Nuno da Cunha dispatched two embassies to Portugal with the news, one by land and another one by sea (v. Morais 1997:41). The construction of the first fortress started immediately and was completed within one year. It also contained the first catholic church on the island. The next year, the Governor Nuno da Cunha assigned a garrison to the fort:[6]

E tendo ja a fortaleza em altura que se podia defender, [...] Ordenou oito centos homens pera ali ficarem de presidio, guarnecendo a fortaleza de artelharia, que tirou dos galeões, e a proveo de muitas monições e mantimentos, deixando alguns navios ordenados pera a serventia da fortaleza. E despedio Isac do Cairo judeu, pera ir por terra ao Reino com cartas a elRey, de como tinha feito aquella fortaleza (Couto 1602)

'And given that the fortress was already high enough to be defended, [...] ordered eight hundred men to remain stationed there, furnishing the fortress with artillery taken from the galleons, and provided it with much ammunition and victuals, leaving behind some ships to serve the fortress. And dispatched the jew Isac of Cairo to travel overland to the Kingdom with letters for the King telling of how he had made that fortress'

However, the area wasn't immediately pacified (see section 3.3 for the significance of this fact). In 1536, the sultan of Gujarat attempted to win back the fort but, during conversations with D. Nuno da Cunha aboard a ship outside Diu, he was murdered[7] (Mir Abu Turab, quoted in Leão 1996:109) and, in 1537, the Portuguese briefly took control

---

[6]While the garrison is described by Couto (1602) as consisting of 800 men, Barros (1615) gives 'novecentos homẽs Portugueses' [nine hundred Portuguese men].

[7]At the time of this episode, according to Barros (1615:book 8, ch.VI), the fort was defended by some 1200 soldiers.

over the city and the customs. As the successor, Bahadur Shah's nephew, was very young, a coalition of three regents temporarily ruled the sultanate of Gujarat. To aid Gujarat and rallied Indian kingdoms in their aim of expelling the Portuguese from the subcontinent, a powerful fleet was prepared in Suez by Turkey's ruler to liberate Diu, which was by then defended by 600 soldiers (cf. Leão 1996:110).

The first siege of Diu started in June 1538 but was lifted after forces from Goa arrived in defence of the Portuguese interests. This created a deep ridge between the Portuguese and the Gujarati rulers, which was settled by an agreement dated 1539 stipulating stern rules of social and economic segregation. Below is an extract of that document, as transcribed in Biker (1881:86-87):

It. Mais hũa parede de largura de quatro covados e mais, da porta da praia do bazar dos que vemdem arroz direito demtro na cidade até a mizquita gramde que estaa em cima do momte e dahy direito até a parede da fortaleza da bamda do mar e a porta que se fizer nesta parede sera aberta todo o dia até hũ quarto da noute e os portugeses e vasalos dos portugeses yram e vyrão e ningem defemderá a eles, e porém ao mouro nom leuarám os portugeses a sua bamda sem licemça do divão, e depois de hũ quarto da noute fecharám a porta. E depois de hũ quarto da noute os portugeses nam ficarám demtro na cidade, e asemtarám nesta porta os oficiaes do mamdouym asy do diuão como dos portugeses pera guardarem os direitos, porém os dos portugeses nom ficarám de noute, e a chaue da porta ficará na mão dos criados do divão e os oficiaes dalfamdega dos portugueses poderám pousar na cidade.

[. . .]

It. toda a remda dalfamdega de gogala e dalfamdega da cidade de dio e asy as remdas das quimtas da jlha sera junto na alfamdega gramde e farám tres partes, as duas partes seraa do divãao e a hua parte seraa dos portugeses.

[. . .]

It. Mais na jlha no lugar que os portugeses tinham feitos dous baluartes e se agora deneficaram na gerra, os quaes nam começarám outra vez de nouo.

'Item. Another wall four or more cubits thick, from the door of the rice market shore inside the city towards the great mosque which lies on top of the hill and from there straight to the wall of the fortress by the sea and the door which will be made in this wall will remain open the entire day until a quarter of the night and the portuguese and their vassals will come and go and nobody will prevent them, and yet the portuguese will not take any moor into their quarters without permission from the *divão*[8], and after a quarter of the night they shall close the door. And after a quarter of the night the portuguese shall not remain in the city, and at this door shall sit officials of the customs, both of the *divão* and the portuguese to defend their rights, yet at night none of the portuguese will remain, and the key to this door will remain in possession of the servants of the *divão* and the portuguese customs officers may remain in the city. [. . .] Item. All profits of the customs of goghla and the customs of the city of diu, as well as the tax collected on the island, will be assembled in the great custom house and divided by three, and two thirds will be for the *divão* and one third for the portuguese. [. . .] Item. Further, on the island, where the portuguese had built two bulwarks now damaged in the war, these shall not be rebuilt.'

---

[8]A local Muslim authority.

The segregation intended by the rulers of Gujarat, and the limitations imposed on Portuguese activity in Diu, were matters of contention. The second siege[9] of the fortress by Gujarati and allied forces took place in 1546 but failed once again, and the local ruler once again insisted on a wall of separation.[10] It was only a few years after these events, in 1554, that the Portuguese wrested complete control of the town beyond the walls of the fort and the entire island following the murder of the sultan Mahmud Shah III; the process by which the captain of Diu fort, D. Diogo de Noronha, assumed the ownership of the island is described in the *Tombo de Diu* (Pais 1592:fl 54). It was only then that the transfer of the colonial structure to the town and the lusitanisation of the island began. By the late 16th century, the territory was firmly under Portuguese control and the property redistributed to benefit the Christian community. The 1592 *Tombo de Diu* (Pais 1592) includes a registry of the rural property (known as *chãos* and *hortas*) outside the city, which reveals that even though the Christians were a minority they (whether Portuguese *casados*, i.e. settlers, or native Christians) owned at least as much land as Hindu and Muslim proprietors combined (see also Matos 1999:22*ff*). The construction of the church in Brancavará (Vanakbara, see Map 1.2), at the extreme western end of the island and an important crossing point onto the mainland, begun in 1630, thereby implying the presence of a significant Christian population in the village - the fortress of Vanakbara was built only much later, in 1770 (Bragança Pereira 1938:185).

3.1.3 Portuguese Diu

With the pacification of Diu, the territory joined the political unit which came to be known as the *Província do Norte* 'Northern Province', a collective designation for the Portuguese-controlled areas stretching from Chaul to Daman, and including Bombay and Bassein.

---

[9]The two sieges of Diu are highly celebrated events in Portuguese lore, and their strong impact in Europe resulted in a flurry of poetic accounts of the heroic feats. These include Corte Real (1574), Andrade (1589), and they also figure prominently in Luís de Camões' *Lusiads*, first published in 1572.

[10]The nature of the conflict between the Portuguese and the local powers is revealed quite clearly in the advice of a Portuguese official, D. Francisco de Lima, to the Governor of India concerning the possibility of peace with the Sultan of Gujarat, in 1548 (transcribed in Biker 1881:128-129):

> deue v. s. de trabalhar pola fazer de maneira que nam pareça ou estê mui craro poder auer rompimento em tempo [. . .] o que está mui certo ser comsentimdo a el Rei de cambaia fazer parede em dio sobre a qual nacêrão as pemdemsas pasadas e foy causa de pôr a jmdia no estado que v. s. vio, porque mui craro está el Rei de cambaia nam querer esta parede senam pera com ela se forteficar comtra nós [. . .] polo que a mim parece que por nhūa cousa se deue de comsemtir na paz com esta parede porque a tenho pola mais certa guerra que se nos pode fazer

> 'your excellency should strive to achieve it [i.e., peace] in such a way that it does not seem or become too clear that there might eventually be a fallout [. . .] which will certainly happen if we allow the king of cambay to build a wall in diu, over which were born the past contentions and which was the cause of putting india in the state that we have seen, for it is very clear that the king of cambay does not want this wall for anything other than to fortify himself against us [. . .] and therefore it seems to me that in peace we should not allow this wall to be built because it seems to me it will certainly lead to war'

Although physically separate, Diu was administratively counted as part of the *Província*, but it was Bassein that emerged as its capital.[11]

The commercial prosperity of Diu was not immediately affected by European rule, fuelled as it was by a dynamic East-Africa/Gujarat route trading gold and ivory (from Africa), textiles and beads (from India) (Vardarjan 1989:372). Among the various communities of Diu, the Parsis, Muslims and Banyans (Vanias) are said to have been responsible for most of the trading activity; as Vardarjan (1989:374) puts it, '[l]ocal communities and indigenous business enterprise [. . .] played a vital role in the prosperity of Diu'. The prosperity of the settlement was recorded by several foreign travellers in the late 16th century, some of whose descriptions are quite eloquent:

> Dui [*sic*] (. . .) is scituate in a little Iland in the Kingdome of *Cambaia*, which is the greatest strength that the *Portingals* have in all the *Indies*, yet a small Citty, but of great trade, because there they laden manye great Shippes for the Strait of *Meca* and *Ormus* with merchandize, and these Ships belong to the *Mores* and Christians, but the *mores* cannot trade neither saile into those seas without the licence of the Viceroy of the king of *Portingale*, otherwise they are taken and made good prises. (English translation of the *Viaggio*, de Federici 1588:f. 5)

> Diu is the strongest town that the Portugales have in those partes. This is but little, but well stored with merchandise for here they lade many great shipes with diverse commodities for the streits of Mecca, for Ormus, and other places. . . the Moores cannot pass except they have a passeport from the Portingales (Fitch 1599)

Notice the three recurrent themes in both Federici and Fitch's descriptions of the settlement, viz. its wealth, the trade links to the Strait of Mecca and Ormuz and also its small size. Bocarro, in his 1634 description of the island, also refers to the 'náos de Mequa' [vessels of Mecca] which took refuge in the port of Diu. The apparent success of the *cartaz* system - which allied commercial entreprise to military violence - is another interesting allusion of these excerpts.

3.1.4   Decline

The decline of Diu as a commercial centre began in the 17th century. It is traditionally associated to the rise of nearby Surat[12] (seem Map 3.1), where the English established a trading post in 1612. The retraction of the traditional trade routes and competition from other operators were also influential in this decline, as mentioned explicitly in Bocarro's 1635 description of Diu:

---

[11]Bassein flourished as an economic centre under the Portuguese, to the extent that Garcia da Orta, in his 1563 treatise on Indian spices, drugs and precious stones (Orta 1563), established that its revenue was 'coisa mais grossa que Diu' [larger than that of Diu].

[12]Surat was the foremost commercial centre for the English East India Company before England received the island of Bombay from Portugal in 1661 and succeeded in diverting much of the trade previously carried out in the Portuguese possessions along the Gulf of Cambay. As Pearson (1976:23) puts it, 'Surat by 1644 was the greatest port in Gujarat (and India)'.

Tem esta fortaleza de Dio alfandiga que rendeo ja cem mil pardaos de mamudes ou larins, que he o mesmo. Não chega a render oje secenta mil, nascido do grande trato e comercio que tem os vaçãos do Mogor de Surrate com a guarda que lhe dão os rebeldes e ingrezes, e de hirem tambem de por Mangalor e Cacha alguas naos sem cartas, as quaes todas levão o mesmo que as naos de Dio, e assy, de forçado, a-de ficar demenuido o que ellas levão. Mas tambem muita parte desta demenuição nasce das tiranias que, em Meca e Suaquem e Adem, fazem os Turcos aos mercadores

'This fortress of Diu has a customs house which used to generate profits of one hundred thousand *pardaos de mamudes* or *larins*[13], which is the same. These days it does not even generate sixty thousand, due to the great commerce of the vassals of the Surat Moghul because of the protection of the rebels and the english, and also because some ships travel to Mangalore and Cacha without permits, all of which carry the same as the ships of Diu and therefore, consequently, diminish what these carry. But a great part of this reduction is also born of the tyrannies which, in Mecca and Suaquem and Aden, the Turks do to the merchants'

As a further reason for the decline of Diu, Bocarro declared that

houve outra causa porque esta fortaleza chegou a tanta diminuição que foi a grande tirania dos capitães para com os moradores assim Portugueses e cristãos da terra como gentios e de qualquer outra lei. (Bocarro 1635)

'there was another reason why this fortress diminished so, which was the great tyranny of the captains towards the dwellers both Portuguese and local Christians as well as gentiles and of any other law.'

In his description, Bocarro refered to the economic decline of Diu but also to stark population decrease, which he attributed to bad governance as well as a particular period of pest and famine. Concerning the size of the town outside the fort, he said

serão oje tres mil fogos, avendo ja sido dez mil. Porem as ditas tiranias dos capitães, ouvidores e mais ministros de Sua Magestade apertarão tanto com elles por lhe trazerem sempre mais e mais, que os forão deminuindo ate este numero, onde tambem a fome e peste do anno de mil seiscentos trinta e dous consumio grão parte desta gente. (Bocarro 1635)

'today it must have some three thousand dwellings, having before been ten thousand. Yet the said tyranny of the captains, officials and other ministers of Your Majesty oppressed them so much in order to receive more and more, that they progressively diminished them to this number, and also the famine and pest of the year one thousand six hundred and thirty consumed a great part of these people.'

This quote is furthermore relevant to the study of the island's social relationships under Portuguese government. Certain documents kept at the *Goa Historical Archives* in Goa attest to the conflicts of interest between a local merchant class and the Portuguese

---

[13]17th-century monetary units. The *pardau* was a gold coin issued by the Portuguese in India inspired on a similar currency of Vijayanagar, and a *larim* was a Middle Eastern silver coin current e.g. in Ormuz.

officials. In a letter dated March 26th 1686[14], the Viceroy Francisco de Távora, count of Alvor, addresses the *Castelão*[15] of Diu under these terms:

> E Por que com aimsolencias e tiranias deque pella mayor parte usarão os capitães desta Praça no tpo de seus governadores digo governos se acha ela arruinada assy pela ausencia que fizerão muitos mercadores como por falta de comercio deste porto, para outros, e dos outros, para este, tereis particular cuidado em ajudar o neg.*o* com todo aquelle favour que for liçito procurando animar os mercadores para que a terra digo elles a continuem com todo o empenho possivel assy para que a terra por este caminho se restetua ao seu antigo ser, como tão bem para que com o mayor rendimento da alfandiga, possa sua Mag.*e* suprir as grandes despesas que aqui continuam.*te* està fazendo e as mais que aodiante puderem ser necessarias.

> 'And because of the insolence and tyranny that the captains of this town mostly practised at the time of their governors i.e. governments it is ruined not only by the absence of many merchants but also the lack of commerce from this port to others and from other ports to this one, you shall be particularly careful to help business with every licit favour, trying to encourage the merchants so that the land i.e. they continue with all possible fervour so that the land may this way recuperate its old self, and also so that His Magesty can with the revenue of the customs compensate for all the great expense continuously made here and the ones that may be necessary in the future.'

While this document is a clear indication of the economic decline of Diu, the very same warnings can be found in letters addressed to other towns in the *Província do Norte*, viz. Daman and Bassein. In Diu, the complaints seem to have originated from the local Banyan traders, who were interested in the development of commerce in Diu but constantly hindered by an abusive administration. One of the attempted solutions to the commercial decline of the city in fact favoured the Banyans in detriment of the colonial structure (v. Section 3.2.2, where the social meaning of this episode will be recuperated).

On the other hand, Diu would become, with its enormous fort, an important penal colony. Several of those judged by the Goan Inquisition during the 17th century were sentenced to confinement or service in the large prison established inside the fort. According to the registry of exilees kept in the *Goa Historical Archives*[16], 28 people were sent from Goa to Diu between 1828 and 1833; the period of their imprisonment varied between 2 and 10 years.

The island's privileged vantage point also meant it was prone to attacks. A Dutch armada is reported to have entered the port of Diu in 1654, but no reference is made to any military engagement. There are references to an Omani-led attack to the fortress, under the context of their opposition of Portuguese presence in the Persian Gulf, which culminated in plundering in 1668 (Leão 1996:40). As the *Província do Norte* came under attack by the Maratha empire from the early 18th century onwards, the walls of Diu fortress (by then already the largest of any Portuguese possession in Asia) were refurbished in 1718 (Morais 1997:125). In 1722, the nearby village of Simbhar (Simbor) was annexed by the Portuguese, but it was lost the following year leaving them with nothing but a military

---

[14] *Goa Historical Archives*, doc. 1376 - 'Livro da Província do Norte'.
[15] The highest political and military authority of the settlement.
[16] *Goa Historical Archives*, doc. 10346 - 'Matrícula dos degredados de Goa'.

post at the site.[17]

In 1736, the dismemberment of the *Província do Norte* began as the Marathas occupied the Portuguese regions of Salsette and Tana, in the vicinity of Bombay. The region of Goa itself as well as Daman were placed under pressure by Maratha forces, but both settlements resisted invasion. In 1739, however, Bassein was taken, which reduced the former *Província do Norte* to Daman, Diu and the factory at Surat (Morais 1997:131). Despite the fact that Diu itself was not under attack by the expanding Marathas, the city was affected by the partial destruction of Portugal's network of influence in northwestern India. Yet, its strategic prominence and fame were not entirely obliterated. Abraham Parsons, a British traveller to the region in the 1770s made the following assessment of the influence of Portugal in Asia:

> while the Portuguese made but little figure in these parts, for, except Goa and the isle of Diu, they had no place of consequence on this side of the Cape of Good Hope.

In the late 18th century, the Portuguese authorities felt the need to reinforce the defences of Diu. In 1773, the fortress of Nagoa was built on the southern shore of the island, outside Diu Town. The French bombardment of the city and the frigate *Real Fidelíssima*, in 1796, is proof of the vulnerability ensuing from the decaying power of the Portuguese in the region. Although the attack lasted for just one day, it revealed that, by this time, the mainstay of the Portuguese possessions of India was the allied power of the British. This attack was interpreted as an attempt on the part of the French to acquire a foothold in the vicinity of Bombay, and the British were concerned to the point of immediately dispatching warships in defence of the Diuese territorial waters (Xavier 1981).

In 1802, British forces disembarked in Diu, Goa and Daman but, faced with the protests of the Portuguese authorities, soon received orders from Bombay to depart; according to Morais (1997), these troops only fully retreat in 1813. In 1837-1838, a political schism resulted in the brief separation of Daman and Diu from the government of Goa. The remainder of the 19th century appears to have been a period of complete stagnation in Diu, which was by then reduced to a quiet, isolated outpost of the *Estado da Índia*, important as a penal colony and little else.[18] Daman, the other remaining possession of the former *Província do Norte*, quickly rose to higher prominence than Diu. The asymmetry between the two is conveniently illustrated by the fact that when the archdiocese of Goa was much later extended to include the rest of Portuguese India (1928), the official designation for the new ecclesiastical post was that of 'Bishop of Goa and Daman', thus leaving out the territory of Diu (Morais 1997:197). There are further indications of Daman's supremacy in the 20th century; whereas a detailed chart of Daman was commissioned in 1914, the same was only commissioned for Diu in 1951 ((Morais 1997:192, 205). On the other hand, the

---

[17]This military post was placed under the supervision of Diu and remained a Portuguese possession until the Indian occupation of Goa, Daman and Diu in 1961. The fort and small stretch of coastline in its vicinity, although detached from Diu, remain part of the Union Territory of Daman, Diu, Dadra and Nagar-Haveli to this day.

[18]Throughout the 19th century, the British authorities repeatedly attempted to secure the donation of Goa, Daman and Diu to the East India Company, which the rulers in Goa always refused. In 1839, they even attempt to buy the territories for 500,000 pounds, to no avail (Morais 1997:163).

(heroic) history of Diu guaranteed a special place for the island within Portuguese India, if anything in emotional and symbolic terms. The *Museu Real Arqueológico da Índia*, the only such institution in Portuguese India with scope over the whole of it, was founded there in 1896 and housed at S. Tomé Church. Seven years later, a commission was created to preserve the archaeological and historical monuments of Diu and, in 1906, the bones of the heroes of the 16th century sieges were solemnly transferred to the Chapel of Sant'Iago inside the fort (Morais 1997:189).

3.1.5   Indian integration

With the establishment of the Indian Union, following the independence of 'British India' in 1947, there arose a conscience in the new nation that the Portuguese colonies of the subcontinent should be integrated. The government of India attempted to negotiate the transition of power, but the Portuguese authorities adamantly refused to consider the possibility. The diplomatic tension between Nehru's India and Salazar's Portugal mounted, leading to the suppression of the Indian diplomatic representation in Lisbon and the blockade of Goa, Daman and Diu (1953), the Indian occupation of Dadra and Nagar-Haveli (July 22nd, 1954), and the expulsion of Portuguese diplomats from New Delhi (1955). On December 18th, 1961, the Indian armed forces simultaneously attacked and occupied ill-defended Diu, Daman and Goa in a military action known as 'Operation Vijay'.

Following integration, a Union Territory was created comprising the former Portuguese possessions. Later, in 1987, Goa achieved statehood and the Union Territory was reduced to Daman and Diu, Dadra and Nagar-Haveli (governed from Daman). As such, the territory nowadays enjoys a special status within India, and in particular towards the neighbouring state of Gujarat. The island's relative prosperity at present is said to owe much to this special position. The sale of alcohol to visitors from neighbouring Gujarat (a *dry state*, where alcohol sale is severely restricted) is a non-negligible source of income for Diu, as is a budding (mostly national) tourist industry.

3.2   Social structure and relations

Even though direct references to the social relations in Diu are scarce, there is somewhat detailed information on the Portuguese policies of colonisation in Asia as well as the groups composing the society of the *Estado da Índia*. As such, it is convenient to start by characterising the history of social structure and relations in Portuguese India as a whole, before discussing to what extent Diuese society corroborates or falls outside the general patterns. In this section, *Estado da Índia* refers to the Indian subcontinent alone, as opposed to the original meaning of the term that included East Africa and the whole of Asia.

3.2.1 Social structure of the *Estado da Índia*

As often pointed out concerning Portuguese colonial expansion, intermarriage was in some areas an official policy.[19] India was one such area, where the colonial rulers decided on the advantages of creating a Eurasian society as a mainstay for political domination. European women were scarce in the region (see below for a more detailed account), which left as an only option to encourage intermarriage between the Portuguese settlers and local women. Afonso de Albuquerque, the first Governor of Portuguese India, explicitly encouraged his men to marry 'algũas Mouras, mulheres alvas e de bom parecer' [some Muslim women, fair and good-looking women] converted to Christianity, as opposed to the darker women of the Malabar coast (v. Boxer 1963:64). To further strengthen the appeal for the creation of a Eurasian community in the 16th century, a Portuguese soldier who married in India was allowed to leave the military career and settle. In 1585, the Franciscan friar Fr. Gaspar de Lisboa attested to the existence of several such families when he wrote

> that in East India there are many generations of Gentiles [i.e. Hindus] who in the course of time adopted the sect of the Moors [i.e. Islam], from which generations descend on the maternal side many sons of India born here whose fathers, even though honourable Portuguese, married in these parts with Christian women of the land whose grandparents and great-parents were of those generations, that is, were originally Gentiles who had become Moors (Fr. Gaspar de Lisboa, letter dated December 14th 1585, quoted in Boxer 1963:42)

As early as 1524, Portuguese men were claimed to be 'todos ou a mor parte, casados com Negras que levam à igreja em cabello muy humtado' (letter of D. Henrique de Menezes dated October 27th 1524, quoted in Boxer 1963:65), i.e. 'all or almost all married to Negresses[20] whom they take to church with very ointed hair'.

Some European families did settle in India, but it is claimed that 'there would seldom be more than a dozen or so of the latter [i.e. women] in a ship which might have six or eight hundred men. Very few married women went out with (or to rejoin) their husbands, and most of the relatively few women who made the tedious and difficult India voyage were the *Órfãs del Rei* 'Orphans of the King" (Boxer 1984:50). The *Órfãs del Rei* were young orphan girls from Portugal sent out to India with government positions as dowries for the men who would marry them. All in all, however, the contribution of entirely European families to the *Estado da Índia* would have been very reduced, even because, as written in 1687, 'ainda hoje he rarissimo o parto de mulher Portugueza, em que não morra, a May, e a criatura' [even today is it still very rare that the labour of a Portuguese woman does not result in the death of the Mother and the offspring] (Padre Fernão de Queirós S. J., quoted in Boxer 1984:51).

There is ample evidence that the mixed-blood community quickly outgrew the European population in the *Estado da Índia*, so much so that an early 17th century writer

---

[19]One caveat is in order, concerning the reconstruction of social relations in the *Estado da Índia*. Many archival documents register official opinions and intentions, but there may be discrepancies between political (declared) intention and actual practice. As such, the very documents which provide priceless information for a reconstruction of social relations in early Portuguese India must be approached critically.

[20]Within the context of 16th-century Portuguese sources on India, *negras* 'negresses' refers either to the local women or African women.

noted that most Portuguese in India had 'more relatives in Gujerat than in Trás-os-Montes' (Diogo do Couto, 1610, quoted in Boxer 1963:78). This community was characterised as a group apart (for their social position, v. section 3.2.2) both in India and abroad.[21] Despite the fast rise of a mixed-blood community [in Goa], Boxer (1963:79) describes its later decline in both status and number, so that in the 20th century they made up a very small fraction of the Goan population.

Society in Portuguese India was highly stratified. As early as 1580, the Jesuit Alessandro Valignano set up a clearly-defined typology of the population in which racial extraction is of paramount importance (see Boxer 1963:62-63). Table 3.2 summarises the structure described by Valignano.

Description	Designation
European-born Portuguese	Reinol
Indian-born of entirely European descent	–
Child of a European man and a Eurasian woman	Castiço
Child of a European man and an Indian woman	Mestiço
Indians	–

Table 3.2: Social structure of the Estado da Índia, 16th-century, apud Valignano

In some written sources, the term reinol refers only to a European who eventually returned to Europe, as opposed to those who settled in India, known as casados (Morais 1997:110). The last category in Table 3.2 is extremely broad, standing for a complex slice of the population made up of several religious groups (Hindus, Muslims, Christians, Parsis,...) as well as different castes and geographical provenances. In addition, there is ample evidence that, from very early on, Europeans other than Portuguese settled in the Estado da Índia. The Archives of the Inquisition of Goa mention several of these; the following process is transcribed as illustration:

[April 4th, 1650]

Luis Fino ou da Criu fraçes de Nação, nat da Cidade de Paris, e residente nesta de Goa por se passar a seita de Calvino (Arquivo Nacional da Torre do Tombo; Conselho Geral do Santo Ofício [Inquisição de Goa], nº33 nº1)

---

[21]One of the missionaries martyrised in Japan in 1597 was an Indo-Portuguese from Bassein (Morais 1997:79); according to a 1686 census reported in Chaudenson (2001: 99), 12 out of 36 families headed by a Frenchman in Bourbon (Réunion) had as a wife 'a "Portuguese from the Indies", i.e., an "Indo-Portuguese, of mixed blood"'. Note the following 1710 description of a Bourbonnais family, in which the wife's name includes both a clear allusion to her provenance and a Portuguese surname (Pereira):

Louis Caron is a Lower Breton aged 68, a notorious and severe drunk ... who, apart from the liquor, is a very honest man, though without education. ... His wife is Monique Péreire Indienne [i.e. Indian], who is conceited, like all the women of that country, though she is unsophisticated and uneducated. (quoted in Chaudenson 2001: 68)

'Luis Fino or da Criu french by Nation, born in the City of Paris, and dwelling in this of Goa [trialed] for adopting the Calvinist sect'

It is important to understand that the military forces which established Portuguese domination along the coasts of the Indian Ocean were incredibly cosmopolitan.[22] In this respect, Pissarra (2002:35) is particularly explicit. He says, concerning the Portuguese forces stationed at Chaul by the time of the 1508 Egyptian-led attack:

A armada portuguesa é uma babel onde se fala uma trintena de idiomas, com destaque para as línguas ibéricas, para o alemão e para o flamengo. Bem representadas estão também as línguas locais e africanas; e ainda o francês e o inglês. Fora estes grupos maioritários, servem a bordo genoveses, florentinos, gregos e albaneses; uma legião de mercenários e escravos com que D. Manuel preenche a crónica falta de gente.

'The Portuguese armada is a babel in which some thirty languages are spoken, with prominence of the iberian languages, german and flemish. The local and african languages are also well represented; and also french and english. Apart from these majority groups, we find genoese, florentines, greeks and albanians serving aboard; a legion of mercenaries and slaves with which king D. Manuel attempts to resolve the chronicle lack of manpower.'

The presence of Indians in ship crews and military forces is not neglectable, even as early as this. The groups most often named in Portuguese sources are the 'Malavares',i.e. inhabitants of the coast of Malabar (roughly correponding to modern-day Kerala), and later also the 'Canarins', i.e. inhabitants of the Kanara and Konkan coastal stretch where Goa is located. In the 1509 Battle of Diu, the 1600-strong Portuguese expeditionary force assembled at Cananor is said to have included 400 *malabares*. Later, the 1521 armada aimed at conquering Diu included 3600 soldiers, 1450 of which are counted as Portuguese and over 2000 as 'malabares e canarins' (Araújo 2008:128). The chronicler João de Barros provides the following description of the 1530 armada sent to attempt the conquest of Diu:

Nesta ilha de Bombaim se fez resenha geral da gente que ia na armada, & acharãose tres mil & quinhentos & sesenta & tantos homẽs de peleja, contando os Capitães, mil & quatrocentos & cinquoenta & tantos homẽs do mar Portugueses com os Pilotos & Mestres, dous mil & tantos Malavares & Canarijs de Goa, oito mil escravos homẽs que pudião pelejar, quatro mil marinheiros da terra q̃ remavão, & mais de oitocentos mareantes dos juncos. (Barros 1615:book 4, ch. XII)

'The general count of the people in the armada was done on this island of Bombay, finding some three thousand five hundred men of arms, including the Captains, some one thousand four hundred Portuguese seamen including the pilots and shipmasters, some two thousand *Malavares* and Goan *Canarijs*, eight thousand male slaves who could fight, four thousand local seamen who rowed, and more than eight hundred seamen of the junks.'

---

[22]When Barros (1615:book 8, ch.VI) describes the episode of the murder of sultan Bahadur, he makes explicit reference to 'hum bombardeiro dos nossos Framengo' [one of our bombardiers, a Flemish]), a man punished for attempting to steal from an inhabitant of Diu.

Given the involvement of all these communities in the affairs of the *Estado*, it is clear that a typology such as Valignano's (Table 3.2) fails to account for the heterogeneous make-up of the population of Portuguese India. One factor that contributed for such heterogeneity is that of the relative ease of movement between the vast territories of the *Estado da Índia*. In this case, the wider meaning of the expression is warranted as population movement is recorded not only between settlements on the Indian subcontinent but also involving other areas of Asia and Africa. The records of the Inquisition of Goa are a convenient source of information regarding these migrations, as both the birthplace and residence of those involved in a process are registered. The following entries exemplify movements within the Indian territories of the *Estado* as well as from India towards a different region:

[April 3rd/4th, 1650]

'Miguel [?], ou dos Anjos, casta guzarate, teçelão, nat da fortaleza de Dio, e morador em Bandorá, penitenciado que foi no auto de 1641 [...]' (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°33 n°1)

'Miguel [?], or dos Anjos, caste gujarati, weaver, born in the fortress of Diu, and dwelling in Bandorá[23], formerly sentenced in 1641'

[October 16th, 1695]

'Fricamo Gentio Guzarate nr.al, e caz.o na fortz.a de Dio, m.or em Mossambique.' (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°33 n°6)

'Fricamo Gujarati Gentile born, and married in the fortress of Diu, dwelling in Mozambique'

African slaves were also brought into the *Estado da Índia*. Descriptions of Goa in the 17th century speak of an incredible abundance of slaves, but this city (the main slave trading emporium of Portuguese Asia) was clearly an outlier among the territories of the *Estado*. Pinto (1992:26) quotes Gemelli-Careri's report of his visit to Goa in 1695, in which he says '[t]here are also an abundance of Cafres and Blacks; for there are Portuguese that keep thirty or forty, and the least six or twelve'. Similarly high figures are also provided in sources pertaining to the religious orders in Goa, the most impressive of which is a complaint by the residents of the Convent of Santa Mónica in Old Goa (unknown date) that their 120 slaves were insufficient, to which is added that even single individuals could have 'fifteen or twenty female slaves, or 26 women and girls, while a *juiz ordinario* or a *desembargador* held 85 female slaves... and some rich ladies over three hundred' (Pinto 1992:27).

According to certain accounts, slaves could be found even in minor Portuguese settlements; Bocarro (1635) wrote, concerning Agaçaim (in the vicinity of Bassein)

a um quarto de légua dela, está uma povoação de uma rua somente, lançada de norte a sul, que tem trinta moradores portugueses, [...] com poucos escravos

---

[23]The former possession of Bandorá, presently called Bandra, is nowadays an area in central Mumbai.

	Male	Female	Birthplace	Age
1	X		Africa	13
2	X		Daman	8
3	X		Daman	6
4	X		Daman	4
5		X	Africa	35
6		X	Daman	1
7		X	Daman	14
8		X	Africa	68
9		X	Africa	35
10		X	Africa	48
11	X		Daman	20
12		X	Daman	14
13		X	Daman	16
14	X		Daman	8
15	X		Africa	75
16		X	Africa	60
17	X		Daman	13
18	X		Africa	14
19		X	Africa	30
20	X		Daman	14
21	X		Daman	5
22	X		Daman	3
23	X		Daman	5 months
24	X		Africa	60
25		X	Africa	35
26	X		Daman	3
27	X		Africa	28
28		X	Africa	60
29		X	Daman	20
30		X	Daman	18
31		X	Daman	38
32		X	Africa	25
33		X	Daman	15
34	X		Daman	8

Table 3.3: Registry of slaves in Daman, 1855

'a quarter of a league away from it, there is a village with one street only, drawn from north to south, which has thirty portuguese dwellers, [...] with few slaves'

While some of Bocarro's claims are rather general, he did advance an average of slaves per colonial household while discussing the city of Bassein. As a caveat, it must be stressed that this was a large and prosperous city, and the figure here may not be extensible to the whole of the *Estado da Índia*. He wrote

[o]s casados que haverá nesta cidade, brancos serão quatrocentos os mais deles fidalgos, com pretos cristãos virão fazer seiscentos e todos terão uns pelos outros três escravos cada um. (Bocarro 1635)

'the *casados* dwelling in this city must be four hundred whites, most of them noblemen, and with the christian blacks it must come six hundred, and all of them must have in average three slaves each'

Slavery was abolished in Portugal in 1836, but there are slave registries in Portuguese India dated 1855 for Daman and Diu. The registry for Diu[24] was set up but never filled in. On the other hand, that of Daman[25] is complete and gives concrete figures on the slave population of the city in the mid-19th c. The information contained in the manuscript is given in Table 3.3.

The owner of each slave is also indicated in the registry, revealing that they were distributed over just 12 households in groups never larger than 6; half of the households counted only 1 or 2 slaves. The distribution of slaves per household is given in Table 3.4.

Master nr.	Nr. of slaves
1	1
2	6
3	5
4	2
5	3
6	1
7	5
8	1
9	2
10	1
11	3
12	4
TOTAL 12	34

Table 3.4: Distribution of slaves per household, 1855

It is also interesting to notice that 46.5% of the slaves were locally-born, all of them under the age of 38. They were probably the children of the African-born slaves registered.

---

[24] *Goa Historical Archives*, doc. 2981 - Registo dos Escravos da Cidade de Diu, 1855.
[25] *Goa Historical Archives*, doc. 2979 - Registo dos Escravos da Cidade de Damão, 1855.

Another implication, not without consequence, is that children born to slave parents were considered slaves. This is also clear from certain earlier descriptions of Portuguese settlements in India. A traveller in Goa, Mandelslo, reported in 1638-1639 (quoted in Chauhan 1995:230):

> Most of the Portuguese have many slaves of both sexes whom they employ not only about their persons but also upon any other business they are capable of, for what they get, comes into the master. Whence it comes that handsome wenches are sought after, to be employed in the selling of fruits and such commodities as the Portuguese send to market, to the end their beauty might draw in customers. Their keeping as to diet stands them in very little. The children born between slaves belong to the master, unless the father will redeem them within eight or ten days after they are born.

Additionally, Thevenot's late 18th-century description of Daman states that 'the Portuguese have slaves there of both sexes, which work and procreate only for their Masters, to whom the Children belong, to be disposed of at pleasure' (quoted in Pinto 1992:28).

According to Boxer (1963:55-57), although slave trade across the Indian Ocean never approached the trans-Atlantic magnitude, it flourished in the 18th century and included both male and female slaves. In the early 19th century, 'the common term throughout the East for an African slave was a "Mosambiquer"' (Boxer 1963:56). The census reported in Table 3.3 indicates that the African-born slaves were originally from the *Rios de Cuama* region, which refers to the Zambezi river delta in Mozambique. It is clear that most of the slaves sold in Asia were therefore from the general area corresponding to modern-day Mozambique, and this is confirmed, albeit relatively late, in a decree issued by the Portuguese authorities in 1816 ruling that the only slaving vessels allowed should be those 'que se destinarem a fazer o Commercio de Escravos nos pórtos da Cósta Oriental de Africa, comprehendidos entre o 10.mo e 25.o graus de Latitude Austral' [destined to trade slaves in the ports of the Eastern Coast of Africa between the 10th and 25th degrees of latitude south].[26]

The Archives of the Inquisition of Goa include several processes involving *cafres* (from the Arabic (*kaffir*) 'infidel'), a term common among the Portuguese to refer to a black person:

> [December 3rd, 1651]
>
> Jorge, em gentio Hiamata cafre natural dos rios de Cuama, e morador nesta cidade por arrenegar da nossa santa fée, e blasfemar contra a pureza de nossa senhora, passandose a seita de Calvino e a de Mafamede em terra de mouros. (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°33 n°1)
>
> 'Jorge, Hiamata in gentile, born at the rivers of Cuama, and dwelling in this city [Goa] for denying our holy faith and say blasphemy against the purity of our lady, adopting the sect of Calvin and that of Mohammed in a moorish land.'

---

[26] *Arquivo Nacional da Torre do Tombo* - Junta do Comércio, Maço 62, Caixa 204. The latitudes mentioned limited the Portuguese possessions in the East coast of Africa and correspond quite closely to the northern and southern borders of modern Mozambique.

This brief survey illustrated the heterogeneity of the population residing in Portuguese India throughout its history. The following section explains the social relations which obtained in the territory.

3.2.2   Social relations in the *Estado da Índia*

In Portuguese India, the European rulers imposed a hierarchy they themselves headed, even though many of the settlers were people of low rank in Portugal. Many of the first colonialists were soldiers or servants of the ships who were encouraged to marry and settle in India. A ship bound for India 'with a crew of 120 men and boys usually carried at least 400 or 500 soldiers sent out for service in the East'; whereas during the 16th century most of these were volunteers and a fraction of criminals or convicts, by the following century 'the majority of the men who were sent to India were not veteran soldiers, but raw recruits taken from the streets and the plough, or convicted criminals collected from the jails and lock-ups' (Boxer 1984:49-50). The social extraction of many among the settlers was the cause of concern or contempt on the part of some of the noblest or most educated among the Europeans. When Ormuz was lost in battle in 1622, for instance, the priest Manuel Severim de Faria was quick to attribute the disaster to the low 'quality' of the soldiers, described as men coming from prison and with no faith in God (Morais 1997:93).

On an earlier stage of Portuguese domination, condemnation of the ungoverned licentiousness of the European men was frequent, in particular on the part of the Christian authorities.[27] Besides the 'recognised' offspring of mixed descent, part of the mixed blood community must have consisted of the children of illicit relationships. It is not clear whether or not this fact weighed in the Europeans' opinion of the Luso-Asian community, but even as this grew in number and identity, they were not always well-received into the ranks of the colonialists. In 1580, Valignano advised against taking *mestiços* and *castiços* into religious orders on account of the European's lack of esteem for them. He also claimed that Indians were unsuited for admittance into the Jesuit ranks

> both because all these dusky races are very stupid and vicious, and of the basest spirits, and likewise because the Portuguese treat them with the greatest contempt, and even among the inhabitants of the country they are little esteemed in comparison with the Portuguese. As for the *mestiços* and *castiços*, we should receive either few or none at all; especially with regard to the *mestiços*, since the more native blood they have, the more they resemble the Indians and the less they are esteemed by the Portuguese (Valignano 1580, quoted in Boxer 1963:62-63)

Western though Valignano's position may be, his remarks present a hierarchy of prestige headed by the Portuguese, followed by the *castiços*, then the *mestiços* and finally the Indians. It is also interesting to notice the claim that this structure was valid from the perspective of the local population. Although Valignano may not have been particularly

---

[27]Father Lancilotto, as early as 1550 complained that many of the settlers had the habit of sleeping with several of their female slaves, and gave Malacca as an example (quoted in Boxer 1963:60-61).

qualified to make such an assertion, the claim ties in very well with the structure of most postcolonial societies.

In 1585 Fr. Gaspar de Lisboa wrote much more favourably about the people of mixed-descent and about how numerous they were. He said, concerning the *mestiços* and *castiços* whose ancestry included Muslims, that 'this is so common here in these parts, that it is no reproach whatever to those sons of India, nor to their Portuguese fathers however honourable they may be, nor is it regarded as a bar to any human honour and dignity, nor up to now has it been the cause of any danger to the faith' (quoted in Boxer 1963:63-64).

The pride the people of mixed descent took in their European lineage led to an often noticed rivalry between *mestiços* and Indians. Writing in 1963, Boxer (1963:80-81) still felt this rivalry to persist in Goa, embittered with the increasing 'impoverishment of the former class [the *mestiços*] and the growth of the latter [the *Canarins*, i.e., the locals] in bureaucratic power and influence during the nineteenth century'. The disbandment of the Army in Portuguese India, in 1871, was particularly resented by the *mestiços*, as it had ensured many of them a job and a position of relative superiority.

Throughout the history of the *Estado da Índia*, the temporal and religious powers were often antagonistic. The Jesuits became particularly influential in the East ever since Francis Xavier began his missionary work in Asia in 1542. According to Newitt (2005:132), the Jesuits' means of achieving power in the *Estado da Índia* had a lot to do with their ability to provide education for the ruling classes. It comes as no surprise that their work in India specifically involved opening colleges in all 'the major centres of Portuguese power in India, starting with Goa and then rapidly extending to the Província do Norte'.

In addition, the views of the temporal and religious spheres of power concerning the local population were far from homogeneous. The sources available for a description of the social relations in the *Estado da Índia* suggest the treatment of the Indians or Eurasians was hardly ever linear. The myth of the absence of colour-bar in Portuguese expansion must be taken with great scepticism, as advised by several historians (Boxer 1963:56), though it may partly apply to (or be based on) the issue of intermarriage when compared to the colonial histories of other European powers. It is not easy to grasp what the overall attitude towards people of mixed parentage, locals or slaves was throughout the *Estado da Índia*; it is likewise difficult to abstract a single Portuguese take on racial (in)equality. Newitt puts these issues in perspective by pointing out the heterogeneity and, to some extent, improvising nature of the Portuguese enterprise in Asia. He writes that

> just as there were two, partly independent forces creating Portuguese expansion - official enterprises and the unofficial diaspora - so conflicting and interacting attitudes to identity emerged. One upheld the primacy of white, European-born and Old Christian Portuguese [...] while the other recognised the Portuguese identity of all those who converted to Christianity, adopted certain symbols of Portugueseness (like the wearing of a hat) or who could claim descent from a Portuguese. Between these polar opposites were many intermediate positions, and issues of race, religion and identity remained in a state of permanent dialectic with one another. (Newitt 2005: 257)

Discrepancy between intention and practice on a political level is also quite obvious. In many instances, it is possible to recognise official attempts at reconciling colonial domination with the integration of the local population and traditions. In 1557, for example,

native Goans were granted the same rights as the Portuguese, a proclamation soon extended to the whole of Portuguese India (Morais 1997:57, 66). At a later stage, in 1761, positive discrimination laws were passed in order to promote the access of Indians to local positions (Morais 1997:139). In the 18th century, the Marquis of Pombal tried to enforce a lay policy of racial equality, also promoting respect for native customs and traditions and making it a criminal offence to use insulting terms directed at Indians; nonetheless, while this enlightened edict was promulgated in 1761, it was only made public in Goa in 1774. Boxer (1963:74) interprets the contrast between the swift application of all other laws passed by the Marquis and this particular instance as evidence of 'how deeply the feeling of racial superiority was implanted in the Portuguese colonial authorities'. Whereas Indians were present in the administration of the *Estado da Índia* by the 19th century, there were strong protests when two of them were nominated in 1822 to participate in the *Cortes* (legislative assemblies) in Portugal.

The accusations of abuse on the part of the administration also come from very early on. Writings by authors such as Gaspar Correia (mid-16th century) and in particular Diogo do Couto with his *Diálogo do Soldado Prático* (circa 1571) are very severe in denouncing corruption in Portuguese India. An interesting case is that of the accusations against the viceroy Conde de Ega, which resulted in his arrest for power abuse; he was posthumously proven innocent (Morais 1997:140, 141), but the episode is indicative of the attempts at controlling corruption in colonial administration. This type of abuse was a common cause of dissatisfaction among the inhabitants of Portuguese India; in reality, Portuguese expansion in Asia saw the explosion of several rebellions (e.g. Batticaloa in 1568, Barcelor in 1582, Timor towards the end of the 18th century,...). In 1787, the so-called *Conjuração dos Pintos* shook Goa itself (Morais 1997:146).

Evangelisation was announced as a primary aim of Portuguese colonial expansion from its very onset, and it received a new impulse with the institution of the *Padroado Português*[28] by the Pope (Morais 1997:36). Although it was not the first diocese in India, Goa received the epithet of *Igreja Metropolitana e Primaz das Índias* 'Metropolitan and Primate Church of the Indies' in 1558 and became the centre of Catholic evangelisation for the whole of Asia. The Roman Catholic Church in India itself sported contradictions when it came to the integration of the local population. On the one hand, the religious structure fulfilled a social role, conveniently illustrated by the nomination of a *Pai dos Cristãos* 'Father of the Christians' responsible for preventing abuse against Indians in Goan prisons in 1565 or the development of educational infrastructures; on the other hand, the Catholic Church strongly resisted accepting Indians or people of mixed descent into important ecclesiastical positions until, by command of Rome and faced with a lack of European vocations, their

---

[28]The Portuguese Patronage was first delineated in 1442. This papal attribution made Portugal responsible for all religious activities in the territories it reached during navigation and exploration, from the Arabian Peninsula to Japan (see Araújo 2008:12*ff*), including the nomination of church officials and administration of its patrimony (Sá 2004:15). Later on, since the foundation in Rome of the Congregation for the Propagation of the Faith (*Sacra Congregatio de Propaganda Fide*) in 1662, conflicts ensued between the Holy See and Portugal concerning the exact jurisdiction of the *Padroado*. Despite successive limitations and negotiations, Portugal retained the right to exercise religious power in some of the areas over which it had lost political and economic control - such was the case of modern-day Maharasthra, Karnataka, Kerala and Tamil Nadu, whose dioceses remained under the jurisdiction of the archbishop of Goa until 1886 (Mangalore, Bombay, Quilon and Madurai) and even later (Cochin and Meliapor).

position had to change in the 18th century (Boxer 1963:65-69).

The criticism that the Church was ready to accept baptisms of convenience rather than conversion arose from within its own ranks. The Jesuit missionary Nicolao Lancilotto, writing from Quilon in 1550, said

> since the inhabitants of these countries are very miserable, poor, and cowardly, some were baptized through fear, others through worldly gain, and others for filthy and disgusting reasons which I need not mention. [. . .] Many people come in order to be baptized, and I ask them why they want to become Christians. Some reply because the lord of the land tyrannizes and oppresses them, and others reply that they must become Christians because they have nothing to eat. (Lancilotto 1550, quoted in Boxer 1963:59-60)

Conversion was several times promoted by means of discriminatory legislation (cf. Boxer 1961); in 1581, a law offered tax reduction for a period of 20 years for anyone who converted. Religious dominance became increasingly more repressive, culminating with the establishment of the Inquisition in Goa in 1560. This court lasted until 1812, with a short interruption by order of the Marquis of Pombal, and had jurisdiction over the whole of the *Estado*. The Goan Inquisition attempted to entirely forbid the practice of native religious rituals through an edict released in 1736, but the political power overruled it (at least with respect to the territory of Goa) a few years later. This is not the only evidence of clashes between the intents of the Church and the colonial government. Morais' list of achievements by the Viceroy D. Luís de Ataíde by 1578 include having been able to 'moderar os excessos praticados pelo clero e pelo Santo Ofício, que eram a causa da emigração dos indígenas' [moderate the excesses of the clergy and the Inquisition, which were the cause of indigenous emigration] (Morais 1997:69). The repression of the Inquisition, with its limitations on religious freedom and consequent emigration of the population, is variously singled out as a reason for the decline of the *Estado da Índia* as a whole (Leão 1996:138). The Archives of the Inquisition of Goa do attest to the extension of its influence, as they include cases referring to the Indian territories of the *Estado da Índia* as well as others in East Africa and Asia - various examples are quoted above, and see also 3.2.3.

A surprising result of evangelisation in Goa (but apparently nowhere else) was the emergence of a system of Christian castes as a compromise between Christianity and the indigenous social structures. The Christians were divided over five hierarchically organised castes; reflecting the Hindu tradition, these were headed by the Brahmin, followed by the Chardos, Sudras, Corumbins and Farazes (cf. Boxer 1963:75-76). Intermarriage was not permitted between the different castes and, although the Catholic Brahmin claimed and enjoyed superiority, the Chardos disputed their protagonism.

3.2.3 Focus on Diu

Despite over four centuries of Portuguese rule, a recent anthropological survey of Diu has found it to be 'an extension of Kathiawad cultural zone which influences language, dress patterns, food habits, customs and practices and other cultural elements' (Singh et al.

1994:5).[29] Such continuity with the cultural panorama of the region is not surprising, in particular as the history of the colony suggests a relatively low degree of cultural intervention on the part of the rulers. In fact, although colonial rule was successfully enforced in the territory (and perhaps because this enforcement was particularly challenging, see 3.1.2), the native population retained a high degree of influence.

One of the communities most central throughout the history of Diu are the Banyans (Vania)[30], who were not circumscribed to the Gujarat region but particularly dominant there.[31] The Diuese trade before Portuguese rule and also afterwards depended heavily on the local Banyan traders, who were widely respected as entrepreneurs. Bocarros's account of the prosperity of Diu is very eloquent:

> Um destes gentios tem hum lugar que chamão Capitão dos Baneanes, que serve como seu procurador em todas as materias que lhe são necessarias e com quem se tratão tambem as tocantes a elles. [...] ha entre elles alguns muito ricos e, se forão favorecidos e ajudados, fora esta hũa das populozas cidades do mundo. (Bocarro 1635)

---

[29]*Kathiawad* is an old appelation of the southern area of the peninsula of Gujarat, which has now been replaced by (and included in) the designation of *Saurashtra*, see Map 3.1.

[30]Unlike other communities, the *Banyan* designation is not strictly bound to a religious group. As Vardarjan (1989:373) explains,

> Vanias could be Meshri-Hindu, or Shravak-Jains, the latter being more numerous. Unlike Decanni Jains, those of Gujarat abstained from agriculture. Their prime occupation was, therefore, trade. [...] In 1646 it was reckoned that 30,000 Vanias were residents of Portuguese India for whom the headquarter was located at Diu.

[31]Many Banyans settled in Mozambique even before the end of the 17th century, under the protection of the Jesuits. They were extremely unpopular among most, including the political administration, and their commercial prowess bitterly resented. Their activity was made easier by the royal decree of 1775, which made the trade of Mozambique free for all the inhabitants of Portuguese Asia in an attempt at refreshing trade along the coasts of East Africa.

> tendo consideraçaõ a que os meyos, e differentes administrações, com que até aqui se tem procurado adiantar o commercio de Moçambique, e mais terras da Africa Oriental, sujeitas ao meu Real Dominio, naõ tem sido bastantes a conseguir hum fim taõ importante [...] Hey por bem extinguir a fórma, porque actualmente se faz este Commercio, e administraçaõ, que se tinha concedido ao Conselho da Fazenda do estado da India, e ordenar, que da publicaçaõ deste em diante fique o Commercio sobredito de Moçambique, e dos mais pórtos, e lugares da sua dependencia, livre para todos os moradores de Goa, e das mais partes, e terras da Asia Portugueza (*Arquivo Nacional da Torre do Tombo*, Junta do Comércio, Maço 62, Caixa 203)
>
> 'considering that the means and different administration with which we have hitherto attempted to further the trade of Mozambique, and other lands of East Africa which are under my Royal Domain, have not been sufficient to achieve such important an aim [...] I have decided to extinguish the way in which presently we do this trade and administration, which had been assigned to the Council of the Treasury of the state of India, and to decree that from this publication onwards the abovementioned trade of Mozambique, and further ports and the places under its jurisdiction, be free to all the inhabitants of Goa and further areas and lands of Portuguese Asia'

'One of these gentiles has a position called Capitain of the Banyans, who functions as their representative in all matters necessary to them and with whom to deal in what concerns them. [...] there are among them some who are very rich and, if they were favoured and supported, this would be one of the populous cities of the world.'

One 17th-century dispute recorded in archival material attests to the role of the Diuese Banyans. In 1686, the viceroy Francisco de Távora visited Diu to hear the complaints of the Banyans; his ensuing proclamation, effectively bashing the colonial administration and transferring economic responsibility onto the Banyans, is particularly indicative of the prestige this community enjoyed in the territory:

para evitar as queixas que havia das oppreçoens que os capitães desta praça farião aos mercadores dellas com excessivos emprestimos q lhes pedião para seus comerçios, ordenou que daqui em diante não houvesse capitães nesta praça, e o governo das armas della se encarregarão a hum Castellão, e no tocante as comercio p.a o porto de moss.e e mais partes correria por conta dos dittos Baneanes, os quais para o d.e efeito formarião entre sy sua comp.a em que entrarião os homs de neg.o e mercadores que quisessem (*Goa Historical Archives*, doc. 1376 - Livro da Província do Norte)

'to avoid the complaints there were concerning the oppression of the captains of this town towards its merchants with excessive loans they asked for their commerce, he ordered that from now on there should be no captains in this town, and the government of arms should be given to a *castellão*, and with respect to commerce for the ports of mozambique and other areas, this should be controlled by the said Banyans, who for the effect would form among themselves their own company of which would take part the businessmen and merchants who so wished'

The Viceroy's call was attended; the Diuese Banyans did indeed organise a *Junta de Comércio* (a commercial company) in 1695 and therefore retained their role as the mainstay of the island's economy. A particular section of Diu Town is still known as *Bairro dos Baneanes* 'Banyan Quarter', as recorded in Map 3.3 (1865).

The Diuese population is nowadays overwhelmingly Hindu, and this appears to have been the case in the 16th century even though the ruling predecessor of the Portuguese (the Sultan of Gujarat) was Muslim. Bocarro explains that

Tem alem disto a cidade de Dio dos muros da povoação para dentro, hua grande povoação de gentios, a maior parte casta guzarates, e alguns judeus brancos, e mouros, [...] que serão oje tres mil fogos, avendo ja sido dez mil. (Bocarro 1635)

'The city of Diu has, besides these [i.e., the Christians and their slaves], a great gentile population within the walls, most of whom of gujarati caste, and some white jews, and moors, [...] today it must have some three thousand dwellings, having been ten thousand earlier.'

The Muslim community precedes the arrival of the Portuguese, as testified by the age of the main Mosque, known as *Jami Masjid*. It is striking that, in contrast with the professed antagonism between the Portuguese and Muslims, Diu should retain such an old Mosque
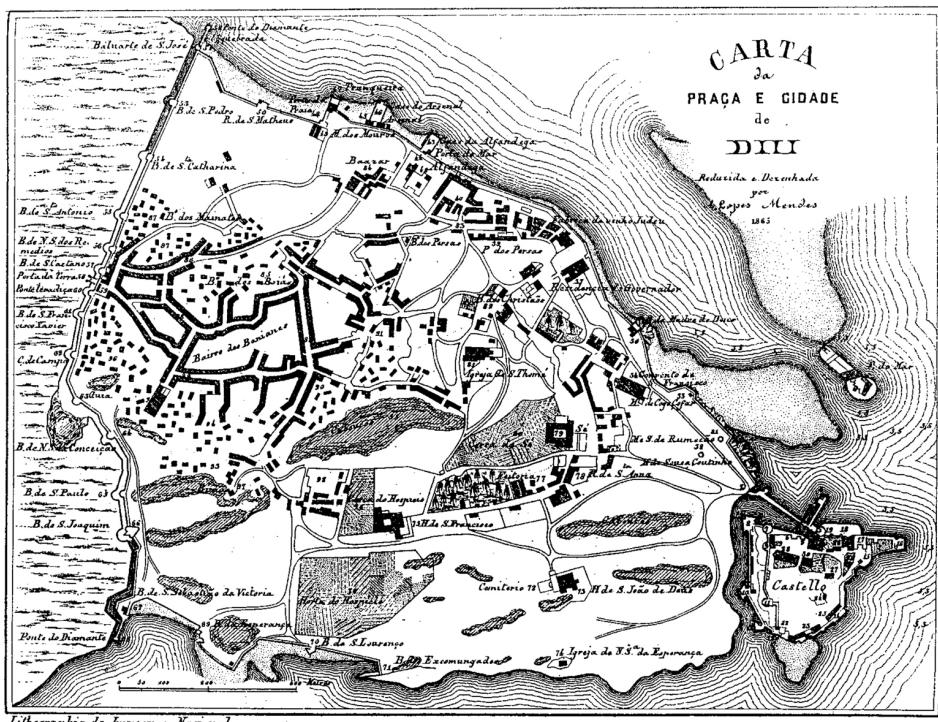
Figure 3.3: Map of Diu Town by Lopes Mendes (1865)

but, as explained in 3.1.2, the conquest of Diu was particularly painstaking and, on the other hand, safeguarding religious integrity was a precondition for the Portuguese to be allowed into Diu. As Shokoohy (2003:2) puts it, '[t]he open hostility against the Muslims does not, however, seem to have been the case in Diu which at the beginning was part of the sultanate of Gujarat, and later close to the territory of the great Mughals, with whom the Portuguese preferred to sustain a friendly relationship, as well as lucrative trade'. In addition, the island's historical trading links included a heavy participation of Muslim traders. Muslim trade in the region was allowed to continue under Portuguese supervision.

The reference to a Jewish community, though explicit in Bocarro, is somewhat problematic. On the one hand, there are are further references to their existence (notice the label *Fabrica de Vinho Judeu*[32] 'Jewish Wine Factory' in Map 3.3), and their presence in as dynamic a trading location as Diu is not at all unlikely. Despite these indications, Shokoohy (2003:2) doubts that there would have been a Jewish community when the Portuguese took Diu. On the other hand, Orta (1563) is very clear when he claims that the Portuguese had the habit of mistakenly calling 'Jews' to the 'Parsis', who were numerous in the kingdom of Cambay.

One cannot be certain of the veracity of the hypotehsis, but if indeed Bocarro's 'Jews' refer to 'Parsis', that would explain why he seems to have failed to notice such an important section of the population. As a matter of fact, the Parsis have a very special connection with the island. According to Parsi tradition (established in the *Kisse-i-Sanjan*, the epic account of the Parsi migration from Persia to India), Diu was the first port of call of the Zoroastrians in India during their flight. They resided there for nearly 20 years before settling in Sanjan across the Gulf of Cambay (see Map 3.1), around 775 AD. Attesting to their long-standing presence in the territory, a complex of abandoned Towers of Silence remains near the southern coastline, and Diu Town also encloses a Zoroastrian Fire Temple. Map 3.3 clearly marks their area of residence as *B^o dos Persas*, i.e. 'Persian [Parsi] Neighbourhood'. The houses in this area of town, which 'compare with the finest Portuguese buildings and residences of the rich Banyan (merchant) Hindus' (Shokoohy 2003:5), still testify to the wealth of this community. The last Parsis are said to have left Diu around 1950, when they ceded the de-consecrated Fire Temple to the diocese. The building nowadays houses the Convent of St. Anne, where nuns run an infants' school.

Map 3.3, as well as other maps even earlier to this one, show a town where different socio-religious groups coexist but occupy different areas. As observed by Dias (2002), after a study of 17th-century cartography,

> No plano puramente urbanístico, parece poder perceber-se que havia duas comunidades separadas, nos extremos, a hindu e a portuguesa, esta dentro da fortaleza, ou castelo, a outra na praça, houve a tentativa de criar um espaço comum, ao longo do rio, virado a Gogola, entre o baluarte de São Mateus e o baluarte da Madre de Deus.

---

[32]*Judeu* appears to have been the designation of a type of wine, and it is therefore unclear whether or not the expression *Fabrica de Vinho Judeu* implies the presence of Jews who produced it locally. Quadros (1899:98) clarifies that '[o] *judeu* é distillado da jagra de canna e das folhas e flôres d'uma planta denominada *daury*, da qual se servem os tintureiros para preparar tintas' [the *judeu* is distilled from cane jaggery and the leaves and flowers of a plant called *daury*, which the dyers use to prepare their dyes].

'On merely urbanistic grounds, one seems to realise there were two separate communities, at the extreme ends, the hindu and the portuguese, this one inside the fort, or castle, and the other one in town, there was an attempt to create a common space, along the river facing Goghla, between the *São Mateus* bulwark and the *Madre de Deus* bulwark.'

The Christian population remains, to this day, mostly concentrated in the oriental section of Diu Town, around the churches of *São Paulo* 'St. Paul's' and *São Tomé* 'St. Thomas''. The history of the settlement (see sections 3.1.2 and 3.1.3) makes it clear that, at first, the Portuguese and their religion were strictly confined to the fort. It was not until after the second siege of Diu that the Christian community began to establish itself outside the fort and toward the existing town. There was a flurry of church-building activity outside the walls at the turn of the 16th century, which probably indicates a period of great development for the Christian community in this area of the city: the church of *São Tomé dos Apóstolos* was consecrated in 1598, the *Colágio do Espírito Santo* (including the church presently known as St. Paul's) was finished in 1606, and the Carmelite Convent of *São José* soon after 1612. The process of physical transfer from the fort onto the town was already very much advanced in 1635, when Bocarro produced the following description:

Mostrão-ce dentro dos muros desta fortaleza grandes ruinas de muitas cazas que nella avia, muy nobres e fermozas, de dous ou tres sobrados, onde antiguamente moravão muitos cazados portuguezes com suas familias, os quaes, pella ma vizinhança que lhe fazião os capitães da fortaleza com seus criados e parentes, largarão as ditas cazas e se paçarão a viver fora, deixando-as cair e chegar aquele estado.

'Inside the walls of this fortress one can see big ruins of many houses which used to be in it, very noble and beautiful, two or three storeys high, where in the past many portuguese *cazados* used to live with their families, which, because of the bad company of the captains of the fortress with their servants and family members, abandoned the said houses and began to live outside, allowing them to fall and reach that disrepair.'

The Christian town seems to have spread right outside the fortress walls but, as this posed considerable difficulties for the defence of the territory, the inspectors sent on a mission to the *Província do Norte* in 1633 ordered, among other things, that 130 houses in the vicinity of the fort should be demolished. This created a physical void of some two kilometres between the fort and the town, which remains to this day.

The survival of the temples of various different cults is rather striking, particularly given the recorded religious intolerance of the Portuguese authorities in other territories of the *Estado da Índia*. It is a well-known fact that, in terms of religious and racial (in)tolerance, Diu constituted something of an exception (see Bragança Pereira 1938:393). In the second half of the 16th century, with the installation of the Inquisition and under Jesuit influence,

the heat was turned on the Hindus and Buddhists in Portuguese Asia, as it had previously been on the Muslims. *With the notable exception of Diu*, wherever else the Portuguese exercised effective power in India and Ceylon, they destroyed the Hindu and Buddhist temples, suppressed the public exercise of all religions other than the Roman Catholic form of Christianity (Boxer 1963:81; my italics)

It is important to realise that the Muslim community was politically dominant in Diu (even though possibly not so in demographic terms) when the Portuguese took control of the territory. The various treaties established with the Sultan of Gujarat aiming at protecting the integrity of the Muslim community and their interests on the island (see section 3.1.2) is in fact in stark contrast with the violent animosity towards Muslims which characterised the actions of the Portuguese in the Indian Ocean. This is quite explicit in the letter sent from the king D. Manuel I to the viceroy D. Francisco de Almeida in 1505. With commercial as well as propagandistic intentions, he ordered, concerning Soffala,

> [. . .] e os dittos mouros catyvares e aos naturaes da terra nam fares dano asy em suas pessoas como em suas fazendas, porque todo queremos que seja guardado, dezendolhe que os ditos mouros que mandamos catyuar e tomar todo ho seu o mandamos asy fazer por serem imiguos da nosa samta fee catholica e com eles teermos contynuadamente guerra (*Regimento* for D. Francisco de Almeida, quoted in Boxer 1963:42)

> 'and you shall capture the said moors, and as for the natives of the land you shall not harm their persons as well as their belongings, telling them that we order the said moors to be captured and deprived of their belongings because they are enemies of our holy catholic faith and we are in continuous war with them'

The relative isolation of Diu with respect to the other Portuguese territories, the difficulty of the process of domination and pacification, as well as its high commercial value and the essential role of various communities for the island's prosperity must not have been alien to the decision to try and promote peaceful coexistence rather than religious antagonism. The catholics of Diu were, however, not exempt from the jurisdiction of the Goan Inquisition. There are several cases moved against inhabitants of Diu recorded in the Inquisition Archives. Some are transcribed here, not only for their relevance in religious terms but also because they provide important information concerning the make-up of the Diuese population by the mid-17th century:

> [March 28th, 1651]
> Mattheus d'Orta mistiço, nat. m.^{or} em Dio, soltr.^{o} filho de Thome da Costa d'Azevedo, por se passar a seita de Mafamede no exterior em terra de mouros sendo prizioneiro. (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], nº32 nº2)

> 'Mattheus d'Orta mestiço, born and dwelling in Diu, single son of Thome da Costa d'Azevedo, for adopting the sect of Mohammed on the outside having been made a prisoner in a Moorish land'

> [April 4th, 1650]
> Diogo Alvres mistiço solteiro filho de Domingos Alves de Carvalho natural de Damão e assist.^{e} na fortaleza de Dio, por culpas de mouro (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], nº33 nº1)

> 'Diogo Alvres single mestiço son of Domingos Alves de Carvalho born in Daman and an assistant at the fortress of Diu, guilty of being a moor'

[March 10th, 1651]

Gaspar do Rosario natural de Chorão morador e cazado em Dio, por se passar a seita de Mafamede no exterior somente em terra de Mouros sendo tomado prisioneiro (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°33 n°1)

'Gaspar do Rosario born in Chorão dwelling and married in Diu, for adopting the sect of Mohammed on the outside only having been made a prisoner in a Moorish land'

[December 3rd, 1651]

Agostinho Ferreira casta pegû natural de Cochim e morador em Dio, por blasfemo, e arrenegar de nossa santa fée, e passar a seita de Mafamede em terra de mouros (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°33 n°1)

'Agostinho Ferreira of the *pegû* caste born in Cochin and dwelling in Diu, for blasphemy, and for denying our holy faith, and adopting the sect of Mohammed in a moorish land'

[March 27th, 1651]

Francisco casta Guzarate, soltr.^{o} f.^{o} de Pays gentios, nat., e m.^{or} em Diu, escravo, pella mesma culpa (*Arquivo Nacional da Torre do Tombo*; Conselho Geral do Santo Ofício [Inquisição de Goa], n°32 n° 2)

'Francisco of Gujarati caste, single son of gentile parents, born, and dwelling in Diu, slave, for the same reason'

The first two cases attest to the existence of *mestiços* in both Diu and Daman. The last quote refers to the local Gujarati population of the territory, while all others show that people born in several locations settled and married in Diu.[33]

Perusing the registry of expenses of the *May de Deus* convent referring to the years 1761 through 1772[34] reveals common offerings to help new converts:

- July 1762: 'Do sustento de m.o més a duas xpãs novas...............04:0:00'
[Half a month's provisions for two new christians..........................04:0:00]

- October 1765: 'De sustento a huma nova xpam..........................03:0:00'
[Provision for one new christian..03:0:00]

- October 1770: 'De sustento a sinco novas xtãs e hũ novo xtão.....010:0:00'
[Provision for 5 new christian women and a new christian man......010:0:00]

- March 1772: 'De esmolla a hũa criança nova xpã........................003:0:00'
[Offering to a new christian child...003:0:00]

---

[33] The fact that colonial mobility, described for the whole of the *Estado da Índia* in section 3.2.2, is seen at play here with respect to Diu must not be downplayed. The movement of populations is likely to have linguistic effects, see section 3.3.

[34] *Goa Historical Archives*, doc. 10332 - 'May de Deus de Diu; despesas'.

The strategies for the propagation of Christianity in Diu were not very different from those observed elsewhere in Portuguese India. The occupation of Diu was relatively late, considering the overall history of the *Estado da Índia*, at a time when religious conversion was already in full swing elsewhere and a (Eurasian) Catholic community in India was already taking form - see section 3.3 for a discussion of the linguistic relevance of this observation. The following section charts, to the extent possible, the development and demographics of the Catholic population in Diu.

3.2.3.1 Evolution of the Catholic population

The Catholics of Diu, unlike those of Goa, are not organised in castes (Moura 1901, Bragança Pereira 1940). Demographic data concerning the Christian and/or Portuguese population of Diu is scarce. The information available is provided here as a contribution towards reconstructing the development of the community which is central to the development and the history of Diu Indo-Portuguese.

As referred in section 3.1.2, the Portuguese who were left in Diu to erect the fortress in 1535 (figures unknown) were supplemented the following year with 800 soldiers. Leão (1996: 110) reports that, during the first siege of Diu (1538), the main fort was defended by 600 of the Portuguese forces, whereas the stronghold in the *Villa dos Rumes*, i.e. Goghla, counted 64 men. The 1635 description of the settlement by Bocarro indicates that

> os casados portuguezes, que vivem oje nesta cidade de fora da fortaleza, são cincoenta e nove, avendo já sido muitos mais: são pobres, pelas ditas causas, mas ainda assim tem huns por outros cincoenta e nove escravos que possão tomar armas, as quaes tem alguns de cabides de lanças muy bastantes pera brigarem.

> 'the portuguese *casados* living in this city today outside the fortress are fifty-nine, although they used to be more numerous: they are poor, for the previously mentioned reasons, but among them still have some fifty-nine slaves capable of taking up arms, some having enough spears to wage war.'

It is convenient to point out that Bocarro's count typically only included adult men who could take up arms in a time of need. However, there are explicit accountd of the presence of Portuguese women residing in the fort even during the bellicose period leading up to the second siege of the city. Several chronicles highlight the pivotal role of women during the first siege, in 1538, when they took it upon themselves to repair the damage caused by enemy attacks and to encourage the soldiers (see e.g. Barros 1954). Several of these women are named and their family status mentioned, which provides a glimpse into the constitution of the feminine population of Diu fort at the time; Barbara Fernandez, for instance, was an elderly widower whose two sons served the fort; Isabel de Veiga, 'filha de hum nobre cidadão de Goa' [daughter of a noble citizen of Goa], was a young woman married to a *reinol* judge in Diu and mother of two; Anna Fernandez was 'de idade velha' [of old age], the wife of a local dignitary.

These women, described as Portuguese, and their families certainly accompanied their husbands and fathers on assignment in the fortress, but there is no indication of their

numbers or the presence of Indian women inside the fort at the time. This reference does make clear, however, that the figures provided by Bocarro must be interpreted; it only refers to *casados*, i.e. European men who had set up residence in Diu, so in order to have an impression of the size of the Catholic community one would have to factor in their families, (most of) their slaves[35] and an unknown number of converted natives. While the earliest Portuguese sources preferably counted the European settlers separately, from the 18th century onwards census data began to take a broader view of the population and to recognise a category of 'Christians'. In 1792, a census of the Christian population provided the data in Table 3.5 - *Sé Matrix* and *S. Thomé* refer to the two parishes of Diu Town (centred around the Churches of St. Paul's and St. Thomas, respectively), and *St. André Ex.* is short for *Santo André Extramuros*, the parish covering the remainder of the island and Goghla (see Map 1.2):

Parishes	Prior of cathedral	Vicars	Benefi- ciaries	Men > 14	Men < 14	Women > 12	Women < 12	Slave men	Slave women	Tot
Sé Matrix	1		2	64	27	114	16	27	43	294
S. Thomé		1		103	23	67	17	15	12	238
St. André Ex.		1		15	12	25	6	3	4	66
TOTAL	1	2	2	182	62	206	39	45	59	598

Table 3.5: Census of the Diuese Catholics - 1792 (adapted from Pinto 1992:31)

Slaves were included in the Catholic population, which means at least part of the slaves in late 18th-century Diu were baptised and possibly integrated to some extent into the Catholic community. In fact, certain members of the Catholic community in modern-day Diu preserve some typically African physical traits, which is easily recognised and confirmed by their peers. The Catholic slave population reported in the 1792 census must not be mistaken for the Siddhi community of Diu, which predates the arrival of the Portuguese and which is nowadays Muslim - see section 3.1.1.

The later census of 1842,[36] given in Table 3.6, surveyed the entire population of the territory. While the overall figure is broken up according to religious groups (*cristãos* 'christians', *gentios* 'gentiles [i.e. hindus]' and *mouros* 'moors [i.e. muslims]'), there is a 274-strong 'racial' category - *negros (também chamados escravos)* 'blacks (also called slaves)' - representing nearly 3% of the whole. It is striking that the 1842 'black' population ascends to more than twice that of fifty years earlier, but in 1792 only the Christians had been counted; in 1842, it is not clear whether the resident (Muslim) Siddhis were included in the *negros* category or the *mouros* section, but the former seems most likely. Otherwise, one would have to assume a slave population much larger than anything glimpsed from previous records and also that only a minority of the slaves employed in Diu was Catholic - which is unlikely given the habit of the Portuguese to baptise their slaves (cf. Pinto 1992).

---

[35] As early as 1592, the *Tombo de Diu* (Pais 1592:fl 67v) mentions the issue of 'negros fogidos', i.e. runaway blacks (presumably slaves).

[36] *Goa Historical Archives*, doc. 2997 - População de Dio.

			1-12	12-25	25-45	45-60	60+
Christians	M	Whites	2	2	6	2	1
		Natives	79	39	76	10	10
	F	Whites	2	2	7		
		Natives	42	40	19	3	5
Gentiles	M		1701	994	1013	124	156
	F		1910	1140	650	185	185
Moors	M		28	37	165	69	19
	F		76	120	80	69	31
Blacks (also called slaves)	M		12	19	81	25	15
	F		16	20	56	19	11

Table 3.6: 1842 Census of Diu

The figures in Table 3.7, referring to the whole population of the territory in 1900, were provided by the governor in his yearly report (Moura 1901:38).

		Catholics	Pagans
Diu island	Town	293	4.276
	From Town to Vanakbara	35	5.757
Goghla		9	3.441
Simbor		24	1.839
SUBTOTAL		361	15.313
TOTAL		15,674	

Table 3.7: 1900 Census of Diu (adapted from Moura 1901)

In this count, *Pagãos* 'Pagans' refers indiscriminately to all non-Christian inhabitants of Diu. As described in section 2.4, the governor goes on to add that, of the 361 Christians, only some 240 were native Diuese, the remainder consisting of Goan soldiers and civil servants on duty in Diu. He further adds that the Catholic population of Diu was decreasing, as there were more deaths than births within the community at the time.

The census of 1921 revealed that, out of a total population of 13,844 people, 12,576 were Hindu, 821 were Muslim, 228 were Catholic, 142 were Jain and only 74 were Parsi. An official statistical publication for the *Estado da Índia*, concerning the year 1957, reveals that in the census count of 1940 there were 306 Catholics for a total population of 19,731, while in 1950 there were 393 Catholics and the population ascended to 21,138 people.

Following integration with India, the overall population of Diu augmented significantly (reaching 39,488 in 1991), while the Catholic community remained somewhat static. Despite high mobility between Daman and Diu (in particular) but also Goa, the Diuese Christian population is presently estimated at around 250 - see section 2.1.

Table 3.8 collates the data presented in a manner easier to interpret. Notice that the focus of demographic information was initially the 'Portuguese' population and only later, from the 18th century onwards, began to contemplate the 'Christian' or 'Catholic' community as a whole.

Year	Portuguese	Christians	Overall pop.
1535		unknown	
1536	1200		
1538	664+		
1635	59 adult men		
1792		598	
1842		347+	9,373
1881			12,636
1887			13,206
1889		c. 300	
1900		361	15,674
1921		228	13,844
1940		306	19,731
1950		393	21,138
1991		c. 250	39,488

Table 3.8: Evolution of the Diuese population, 1535-1991

These scattered indications, along with the historical information explored in section 3.1, permit some general observations concerning the evolution of the Diuese population and, in particular, that of the Christian community. It is clear that the initial thrust of settlement in Diu (from 1535 onwards) brought over a great quantity of soldiers, sent to Diu from the territories already under Portuguese control in India. Recall that the numbers referring to *Portuguese* are likely to refer to men under Portuguese command rather than necessarily Portugal-born individuals; as discussed in 3.2.1, there is concrete evidence that the Portuguese-led naval and military forces were highly cosmopolitan. A steady influx of people from the remainder of the *Estado da Índia* was constant at least until the end of the second siege (1546).

The economic decline of Diu (see section 3.1.4), from the 17th century onwards, was accompanied by a steady decline of the Portuguese population; not long before 1635, an outbreak of cholera took a toll on the Diuese population in general, as mentioned by Bocarro (1635). Apart from this episode, and with the exception of yet another cholera outbreak in the late 19th-century and a short population decline in the early 20th century, the overall population of the territory appears to have increased steadily but slowly (and considerably stepped up after the decolonisation of the island).

Interestingly, the Catholic community did not accompany this growth; there are no precise descriptions of the process by which a local (either Indian or *mestiço*) Catholic community formed in the early days of Portuguese domination, though it is clear that this section of the Diuese Catholics soon outnumbered those described in the censuses as 'white', i.e. Europeans and their descendents (see Table 3.6). Considering how the 16th century was the heyday of Portuguese-controlled missionary activity in India[37], the Catholic community of Diu probably reached its numerical peak in the 17th century, before the beginning of

---

[37]The proclamation of the *Padroado Português* in 1516 is an important benchmark in this process, and 1588 is credited as the year with the most conversions and baptisms as a result of Portuguese missionary activities in the Orient (Morais 1997:75). In Diu, as mentioned before, the late 16th century and early 17th century witnessed an unparalleled flurry of church building and the settlement of religious societies, accompanying the spread of the Christian town outside the walls of the fort.

the settlement's decline. As shown in Table 3.8, we do find hard evidence that from at least the late 18th century, the community has undergone a sustained reduction.

The significance of all these facts for the development of DIP is explored in the following section.

3.3   Sociolinguistic history

The historical indications collated in the previous sections can now be interpreted in order to reconstruct the sociolinguistic setting in which DIP formed and developed. It is fair to concede that, for all the hard evidence concerning post-19th-century Diu, we have remarkably little to support our analysis of the earliest period, which is the most crucial to understand the formation of DIP. Other researchers have faced the same difficulty, to the extent that Bakker (2002:75) admonishes that accounts of creole formation are often 'purely speculative, but they are quoted as fact by other creolists'. My analysis of the formation of DIP is therefore intended as flexible, capable of incorporating further documentary evidence whenever it comes to light. It is clear, at any rate, that the sociolinguistic context of Portuguese in Diu in the early days of the Portuguese settlement was somewhat peculiar when compared to other regions of the *Estado da Índia*. This section highlights some essential elements, tending towards a general scenario of the formation and development of DIP.

Unlike other conquests of the Portuguese in India, the settlement in Diu was clearly not a commercial entreprise, but a military and strategic one. As such, the early history of Portuguese Diu was eminently bellicose and traumatic at least until after 1546. The acquisition of the fort took place relatively late, in 1535, at a moment when other areas of the *Estado* (such as e.g. Goa, Chaul or Cochin) had already been pacified, distinct Asian and Eurasian Catholic communities had already been established[38] and the evangelisation of these territories was in full swing. In fact, Clements (1996) estimates that the Chaul-area variety of Indo-Portuguese acquired native speakers as early as 1620; southern varieties of IP (e.g. Cochin or Cannanore, see Map 1.3) may have been actively spoken even earlier than that. The 1200 soldiers assigned to man the fort of Diu in 1536 were recruited in these areas, and documentary evidence suggests that the defending force, between 1535 and 1546, involved men from various ethnic and linguistic backgrounds (see 3.2.1).

In addition, the physical space of (Indo-)Portuguese was at first highly constrained. The year 1554 (when the Portuguese assumed control over the entire island) is a benchmark date in this respect, marking the beginning of the real domain of Portugal over Diu. Before that, and given the social and physical segregation imposed by sultan Bahadur (see section 3.1.2), the Portuguese population and their vassals were mostly confined to the enormous fort. During this period, large numbers of soldiers under Portuguese command, traders, clergy, officers and their families lived in close proximity, pushed to the extreme during the sieges of 1538 and 1546. Diu Town (outside the fort) was quite hostile to the Portuguese at the time, which is evident from the fact that much of the population welcomed the Turkish

---

[38]Records of offspring of mixed Asian and European descent are available as early as 1516 for Goa (cf. Clements 1996:8).

and Egyptian forces as they laid the first siege to the fort (Barros 1615:book 10, ch. VI).

It would be a gross oversimplification to reduce the linguistic equation eventually resulting in DIP to the contact between Portuguese and Gujarati. The 'fort phase' was probably determinant for the formation of DIP, despite the reduced (if not null) involvement of native, Gujarati-speaking Diuese people. The conditions of coexistence in Diu imposed serious restrictions on missionary activities, and it is therefore a matter of speculation whether a significant Diuese Christian and/or Eurasian community began to form before the second siege at all. In any case, the setting inside the fort appears conducive to the formation of a high-contact variety. Some important sociolinguistic variables which obtained in this phase include:

- *Linguistic heterogeneity*: the presence of various languages and registers, including standard and non-standard Portuguese, the Indian languages of Indian conscripts, various African and European languages and possibly forms of Indo-Portuguese beginning to be nativised across India;

- *Regular social interaction*: enhanced by the physical concentration of the relevant population (the Portuguese and their vassals);

- *Demographic fluctuation*: throughout the 'fort phase', the population of the fort changed considerably, as the structure received regular military reinforcements and also contributed elements for various armadas;

- *Slow nativisation*: the population must have been overwhelmingly adult and male, at times (the sieges) experiencing high mortality rates and low birth rates.

Considering that Portuguese was not an L1 for a large number of the fort's inhabitants (if not the majority), communication may have availed of a restructured Portuguese register, whether locally formed or the general Asian Portuguese Pidgin identified by Clements (2000). However, the 'fort phase' was probably an unlikely setting for its nativisation, unless of course the number of children (particularly of mixed backgrounds) born or raised within the defensive walls were larger than the available records indicate.

The 'town phase', starting in 1554, was certainly more conducive to the nativisation of this restructured variety of Portuguese. Really significant interaction between the settlers and the Diuese only stepped up after the Portuguese took control of the island and the Christian settlement was allowed to spill out of the fort. The conditions were thereafter met for Portuguese *casados* to settle in Diu (including the distribution of land on the island, see 3.1.2) and for a more aggressive christianisation of the territory to begin. In other words, the conditions were met for either the offspring of converted locals and/or the offspring of African-born slaves and/or the offspring of Eurasian couples to acquire a non-standard variety of Portuguese as their first language.

In this phase, the linguistic pool available in Diu was certainly still diverse, including Gujarati and Standard Portuguese, possibly the 'fort phase' pidgin or IP varieties transplanted from different regions. It is clear that most Indo-Portuguese varieties, though physically discrete, did not form or develop in isolation. The evidence provided in 3.2.1

attests to the vitality of an inter-*Estado da Índia* mobility which allowed the population to circulate quite freely. This is an important point, as the constant flux of population can be accountable for the striking similarities between the various Indo-Portuguese varieties.[39] The island's role as a penal colony also brought over a considerable number of inmates from the various areas of the *Estado* (see 3.1.4), in particular after the establishment of the Goan Inquisition in 1560.

Population fluidity was a constant throughout the colonial history of Diu, and marriage patterns favoured the demographic flux. Given the caste principles governing much of Indian society, the tendency of the Christians towards endogamy is hardly surprising. However, considering how reduced the Diuese Christian community has always been (see Table 3.8), endogamic practice often involved marriage to people from further afield. The particularly close link between Diu and Daman (initially the *Província do Norte* in general), which remains to this day, was probably operative from a very early stage, not only on account of their physical proximity but also because the *Província* constituted a single administrative unit. Goa, despite the sheer size of its Christian population, may have been somewhat adverse to the Indo-Portuguese Christians from the north; as reported in Moura (1901:41) '0 *norteiro* não se cruza com os filhos de Goa: a isso se oppõe as castas destes' [The *norteiro* does not cross with the sons of Goa: the castes of the latter are opposed to that].

One of the consequences of the decline of Diu as a commercial hub, from the 17th century onwards, was the reduction of the island's ethnocultural and linguistic heterogeneity. Slave import halted, the presence of Portuguese settlers diminished, and DIP was more clearly circumscribed to the community of Diu-born Catholics. However, normative pressure was not removed from the linguistic picture. From the 19th century onwards, at least, local observers had the impression that change affecting DIP tended towards SP structures. This was attributed, among other things, to a certain development of Portuguese-medium education on the island and the role of Goa(n Portuguese) in Diu and the Diuese families (which however contradicts the observation of Moura above). Jeronymo Quadros is particularly explicit in the following account:

> De resto, em rigor, é agora difficil, se não impossivel, marcar a linha verdadeira a que se circumscrevia o antigo dialecto de Diu, vulgarmente conhecido por *lingua norteira*, tanto devido á escassez de material recolhido e classificado, como pelo progressivo aportuguezamento a que o mesmo dialecto se foi subordinando n'estes ultimos 20 annos. [. . .] uma serie de circumstancias, que acabaram por lhe destruir toda a sua primitiva originalidade, das quaes as principaes são a diminuição da população christã, a quasi fusão d'esta com o elemento goêz, a influencia do ensino official e, em consequencia, o sucessivo desuso das formas glottologicas que eram a sua caracteristica (Quadros 1907:193)

> 'On the other hand, to be precisa, it is now difficult, if not impossible, to draw the true line which circumscribed the ancient dialect of Diu, commonly known as *lingua norteira*, not only because of the paucity of collected and classified material, but also because of the progressive

---

[39]Dalgado was of the same opinion; in order to account for the observed similarities between the various IP varieties, he proposed that one must 'admitir frequente contacto dum com outros e recíproca transfusão parcial, proveniente da constante migração da grande parte dos indivíduos que os falavam' [admit frequent contact between one and the other and partial reciprocal transfusion, resulting from the constant migration of a large part of those who spoke them] (Dalgado 1917:41).

portuguesification the dialect has been subjected to in the past 20 years. [. . .] a number of
circumstances, which ended up destroying all of its primitive originality, of qhich the main
ones are the diminution of the christian population, the near-fusion of it with the goan ele-
ment, the influence of official education and, as a consequence, the ensuing abandonment of
the glottological forms which were characteristic of it'

The perceived 'portuguesification' did not succeed in wiping out DIP, and in fact de-
colonisation appears to have altered the trend somewhat (cf. Tomás 1992:65). Issues con-
cerning the typological allegiance of DIP are dealt with in chapter 9, but it is important
to retain that, as any other living language, DIP has undergone change throughout all of
its history in response to the fluidity of its sociolinguistic context.

Part II

Description

Chapter 4

Descriptive preliminaries

This chapter clarifies some aspects relevant for a correct interpretation of the description of DIP contained in chapters 5 through 8. I will begin by characterising the analysed corpus, both in terms of the methods of collection employed (section 4.1) and its constituency 4.2. Section 4.4 deals with the physical aspect of the example sentences transcribed in this description, both in terms of the adopted orthography and glossing rules.

4.1 Data collection

The data for the present research was collected over 4 field trips to Diu, between 2004 and 2008, through interviews and elicitation sessions involving primarily native speakers of DIP but also non-native speakers, of all generations and from various communities (Hindu, Muslim and Catholic). Apart from the information collected in formal recording sessions, much data was also collected through the observation of, and interaction with, the community of first-laguage DIP speakers. The resulting corpus is described in detail in section 4.2. The mediator language of choice (in particular for elicitation) was English, given that Gujarati was not shared between researcher and collaborators, and (standard) Portuguese was unadvisable due to the possible effect its local prestige might have on the collaborators' output.

The inevitable effects of the *observer's paradox*, exacerbated in Diu by the sociolinguistic constraints described in 2.2.1.3, were contradicted - to the extent possible - in the design of the recording sessions. Therefore, in addition to one-on-one interviews and elicitation sessions, I often aimed at the following settings:

- *Group sessions*: these potentiated free interaction between the various participants and, although more demanding when it came to transcription and phonological analysis, yielded important material which was not necessarily directed at the researcher but at peers. One of the groups which was recurrently interviewed collectively consisted of school children (aged 6 to 12) who attended the (English-medium) Nirmala Mata Catholic School in Diu Town.[1] Other group sessions often involved entire households.

---

[1] These recording sessions were organised during the students' school break, with the kind assistance of

- *Dialogues*: the attempts to record dialogues without the intervention of the researcher were notoriously unsuccessful, given the unease created by the recording gear and the entire situation.

- *Public functions*: whenever possible, I recorded public functions, including the performance of traditional songs in DIP.

- *Mediation*: several recording sessions availed of the assistance of mediators (native speakers of DIP and peers of the interviewees) who translated the questions and requests of the researcher for the interviewee(s), from English into DIP. By shifting the primary interlocutor role onto the mediator, this setting minimised the interference not only of the researcher but also of his perceived expectations and language skills in the production of the interviewee(s).

It became clear very early on that video recording was counterproductive, as it enhanced the disturbance caused by the recording setting. The vast majority of recordings was therefore strictly audio.

The few video recordings were made with a Sony DCR-HC40E video camera, on Mini DV tapes (DVM60PR3). The audio data was recorded on a Sony MZ-NH1 Hi-MD Minidisc, with two external electret condenser microphones: Sony ECM-719 clip-on microphone and Sony ECM-MS907. The audio data was stored as uncompressed *.wav* files with a sampling rate of 44 kHz. At a later moment, it was transcribed using *Transcriber* software and exported to *Toolbox* for interlinear glossing and analysis.

Apart from audio data, information was also collected through various questionnaires and fieldnotes. The following section describes the corpus in detail as well as the criteria determining which data to consider for the present description.

4.2   Corpus

The total audio corpus collected in all field trips to Diu ascended to *circa* 33 hours of continuous speech, divided over 70 recording sessions. This sum total includes all types of recording sessions (see 4.1), including elicitation sessions, and it was therefore necessary to make an educated selection of which material to consider for the linguistic description. The following criteria guided the decisions made concerning which sessions to transcribe and analyse:

- *Fluency*: the material collected through the interviews of non-native speakers of DIP was, for the time being, disregarded, given that the L1 material collected was deemed sufficient and more reliable.

_____

the various school directors.

- *Absence of normative constraints*: collaborators whose speech displayed the phenomena associated with the normative pull of Standard Portuguese (for which, see 2.2.1.3) are not represented in the corpus for analysis.[2] The manifestation of these normative effects was aprehended through observation and comparison of several collaborators' speech. As it turns out, young speakers are in general reliable sources of speech unconstrained by the normative pull of SP, but so are some older speakers.

- *Family history*: only speakers who were born in Diu were considered for the sample. Furthermore, considering the relationship between DIP and Daman IP in the territory and the family effects observed (see 2.2.2), the speech of Diu-born children of Damanese parents was included in the corpus but, whenever necessary, checked for consistency against the speech of other collaborators.

- *Linguistic dexterity*: it is a fact that some collaborators are more capable than others of detaching themselves from the structures of the mediator language, in an elicitation setting. This fact was taken into account when approaching elicited material which, at any rate, is treated as secondary to that obtained from free-flowing speech.

In addition to the audio recordings, this description avails of material elicited in the form of various questionnaires (filled in by 7 collaborators), field notes stemming from the interaction with L1 speakers of DIP, and the elicitation of *circa* 700 vocabulary items made with resort to a *Pictorial Glossary in Indian Languages* published by the Central Institute of Indian Languages, Mysore (Karnataka).

4.3   Phonemic and phonetic transcriptions

For phonemic and phonetic transcriptions, I follow the rules of the *International Phonetic Alphabet* (henceforth IPA) set out by the International Phonetic Association (1999). The orthography used for non-phonemic/phonetic DIP text, on the other hand, differs quite substantially from the IPA; see section 4.4.1 in this chapter for a description of the valid orthographic conventions, and 5.5 for a full motivation of the orthography grounded in the phonemic description of DIP.

Two symbols which are used irregularly are the IPA sign for primary stress - an apostrophe "'" preposed to the stressed syllable - and the IPA sign indicating a syllable boundary - a dot".". These symbols are used, in phonemic as well as phonetic transcription, only when required by the line of argumentation.

---

[2] Aside from the obvious sociolinguistic relevance of the normative pull exerted by SP, data from such speakers is in fact rather interesting in that it highlights a) the areas of DIP grammar perceived by the speakers as divergent from SP, and b) the continuing role of SP in the linguistic context of L1 speakers of DIP. This data was for the moment disregarded so as to achieve a maximally accurate description of unconstrained DIP, but it should in the future be analysed in more detail from a sociolinguistic/variationist perspective.

4.4   Example sentences

The example sentences provided in this description were retrieved from the analysable corpus as defined in 4.2, including both free-speech and elicited material. I resort to elicited data only in the absolute absence of adequate example sentences from free-flowing speech.

Examples are given in italics as a default. Any lexical items not in italics refer to borrowings (e.g. English borrowings, such as *uncle*), and these are given in their original orthography; recurrent loans which are commonly adapted to the phonology of DIP (such as *saykəl*, from Eng. *cycle*, which in DIP occurs with final stress) do not appear in italics and follow the DIP orthography. The orthography of Gujarati, Hindi and Konkani lexemes, on the other hand, is latinised according to the orthography designed for DIP (see 4.4.1) unless a particular author's version is transcribed (in which case this is duly indicated).

The following subsections clarify the orthographic and glossing conventions used in the treatment of DIP data.

4.4.1   Orthography

The orthography used in this description is such that each grapheme (or compound grapheme, as in the case of <ch>) corresponds to a phoneme. The system is fully motivated in 5.5, and therefore I will simply indicate the correspondence and approximate pronunciation here.

- <a> corresponds to /a/: pronounced [a].

- <ɛ> corresponds to /ɛ/: pronounced [ɛ].

- <e> corresponds to /e/: pronounced [e].

- <i> corresponds to /i/: pronounced [i].

- <ɔ> corresponds to /ɔ/: pronounced [ɔ].

- <o> corresponds to /o/: pronounced [o].

- <u> corresponds to /u/: pronounced [u].

- <ə> corresponds to /ə/: pronounced [ə] or [ɐ].

- The application of a tilde on a vowel (<ã>, <ẽ>, <ĩ>, <õ>, <ũ>) results in a nasalised vowel.

- <y> corresponds to /j/: pronounced [j].

- <w> corresponds to /w/: pronounced [w].

- <p> corresponds to /p/: pronounced [p] or [pˀ].

- <b> corresponds to /b/: pronounced [b] or [bˀ].

- <t> corresponds to /t/: pronounced [t] or [tˀ].

- <d> corresponds to /d/: pronounced [d] or [dˀ].

- <k> corresponds to /k/: pronounced [k] or [kˀ].

- <g> corresponds to /g/: pronounced [g] or [gˀ].

- <m> corresponds to /m/: pronounced [m].

- <n> corresponds to /n/: pronounced [n].

- <ŋ> corresponds to /ŋ/: pronounced [ŋ] or [(j)ɰ].

- <r> corresponds to /r/: pronounced [ɾ] or [ʀ].

- <f> corresponds to /f/: pronounced [f].

- <v> corresponds to /v/: usually pronounced [ʋ].

- <s> corresponds to /s/: pronounced [s].

- <z> corresponds to /z/: usually pronounced [z].

- <x> corresponds to /ʃ/: pronounced [ʃ].

- <ch> corresponds to /tʃ/: pronounced [tʃ] or [ʃ].

- <j> corresponds to /dʒ/: pronounced [dʒ] or [ʒ].

- <l> corresponds to /l/: pronounced [l].

4.4.2 Glossing

In this description, glosses are closely modelled on the *Leipzig Glossing Rules* published by the Max-Planck Institute for Evolutionary Anthropology (Leipzig). I will only draw attention onto a few salient glossing practices which may not be easily interpretable. Consider,

as an illustration, example (11):

(11)   *pɛrt d-ikəl*   *ki*   *te*   *ũ*   courthouse.
      near of-DEM$_d$ CMP EXS.NPST one courthouse
      'A courthouse which is near that.'

Words which are italicised in example sentences (but not the glosses, see below) are loanwords, which means the orthography may differ from that defined for DIP text; this is exemplified by the English word *courthouse* in the previous example. Segmentable morphemes are separated with a hyphen "-" in example sentences, as well as in the corresponding gloss; this is the case of *dikəl* 'of that' in (11). When a non-segmentable example word (i.e. in the case of suppletion) requires several gloss elements, they are mediated by a do "."; this is exemplified in (11) with the word *te*, glossed 'EXS.NPST'. Recurrent glosses (usually grammatical) - such as 'DEM$_d$', 'CMP', 'EXS' and 'NST' - are abbreviated. A full list of abbreviations can be found in the opening pages of this volume.

(12)   *dəpəy muyɛ∼muyɛr*   *də Manu.*
      dəpəy woman∼woman of M.
      'Then the women of Manu('s family).'

In reduplicated forms, the constituents are graphically connected with a tilde "∼"; see *muyɛmuyɛr* in (12). Proper names (such as *Manu*) are indicated only by their initial in the gloss. Ocasionally, when required by the argumentation and often to draw attention to form rather than internal structure/semantics, certain words may be left unanalysed in the gloss; in these cases, they are transcribed in italics, such as *dəpəy* in (12).

Whenever an element of an example sentence is enclosed in parentheses, the interpretation is that it is optional. Square brackets may also be used in examples to draw attention to a particular constituent or to indicate phrasal or clausal boundaries, whenever this is required by the argumentation; in these cases, the use and meaning of the square brackets is duly flagged in the text. On the other hand, bracketed elements in the translation line are those which have no correspondence to any (overtly) expressed element in the example sentence but which are required in the translation language (i.e. English). The use of brackets in examples is demonstrated in 13:

(13)   [*a*   *el*] *(de-w).*
      DAT 3s give-PST
      '[I] gave [it] to him.'

Chapter 5

Phonology

The definition of a language's phonology implies the delimitation of a corpus from which to abstract the system. While in certain cases this may be relatively straightforward, in highly multilingual ecologies such as Diu (see chapter 2 for a detailed account) it often turns into a challenge - any solution to which eventually proves somewhat arbitrary. As expected, the speech samples collected among native speakers of DIP indeed contain elements from several distinct etymological sources: Portuguese, Gujarati, (Indian) English, and other South Asian languages. It is clear that the speech I was able to collect is shaped according to the speakers' image of the researcher as an interlocutor as well as his perceived competences and expectations; admittedly, a different interlocutor with a different social and linguistic profile would trigger different linguistic choices on the part of the speakers of DIP. As a result, it is often impossible to produce watertight arguments to include or exclude a particular lexeme from an abstracted 'core' vocabulary of DIP. DIP speakers have at their disposal a vast array of lexemes from various sources, which may or may not be summoned according to the context of each utterance, and the boundaries between hypothetically different 'systems' is blurred as a result of actual practice. This is of course a theoretical problem that arises time and time again in studies of code-switching/mixing and multilingual ecologies (e.g. Myers-Scotton 1993, Mufwene 2001, Mosel 2004), one for which there may not be an uncontroversial solution.

Given these caveats, the criteria adopted in this study so as to define a lexical corpus for phonological analysis is not intended as dogmatic; it is instead driven by practical reasons and based on some general indications of phonological behaviour as well as notions of textual frequency (for which, see section 8.1). The picture that emerges from even a superficial survey of the spoken corpus is that the phonology of a major subsection of the vocabulary - the Portuguese-derived items - differs in significant ways from the remainder of the lexemes employed.[1] It seems warranted to claim that the backbone of the DIP lexicon consists of Portuguese-derived elements, as supported by two observations: a) the overall dominance of Ptg.-derived lexemes in the corpus, and b) the overarching presence of Ptg.-derived items within DIP's closed lexical classes.[2]

---

[1] This is a purely synchronic observation and as such does not imply the absence of phonological convergence throughout the history of DIP. For a detailed analysis of this issue, see section 9.1.2.

[2] When compared to other IP varieties for which there are descriptions (e.g. Korlai IP, Sri Lanka IP, Cannanore IP and Cochin IP), it becomes clear that the functional items of DIP rely strikingly little on

Accepting that the most significant elements for our enterprise are the ones derived from Portuguese is equivalent to positing that the phonological rules shaping them are a sort of phonological 'matrix' which, because of its dominance in the speech of DIP speakers, may exert some adaptive pressure onto words from other sources. As a result of this community's high proficiency in Gujarati and English, that is seldom the case (i.e. Gujarati lexemes need not adapt to what I will consider the *core phonology* of DIP; when it does happen, one can then speak of a more or less permanent borrowing.

The type of reasoning governing my decisions on what phonological evidence to take on board and what not to is exemplified with the discussion of whether the phonemic array of DIP contains retroflex phonemes. There is no question whether or not speakers of DIP are familiar with retroflexes or able to produce them, given that all are proficient in Gujarati from early childhood onwards. A look at the attestations of retroflex sounds in the corpus (which consists of DIP utterances rather than Gujarati or Indian English) reveals that they are restricted to Gujarati and English lexemes. Consider the following examples:

(14)   a.   *el ɛ*   *bẽy bəɖka, el bəɖka.*
            3s COP$_i$.NPST very fearful 3s fearful
            'He gets scared very easily, he gets scared easily.'

       b.   *fuʈbɔl   ku   bɔlibɔl*
            football and volleyball
            'football and volleyball'

In (14a), the word *bəɖka* (from the Gujarati root *bəhka-* 'scare') is selected to describe someone who gets frightened very easily. In (14b) the retroflex plosive in the word *fuʈbɔl* results from a process of phonological adaptation of L2 English to the Indian context by which the alveolar consonants of Standard English tend to be replaced with retroflex consonants (Trudgill and Hannah 2002:130).[3] Faced with these facts, one must decide on the distribution of these items across the lexicon(s) available to speakers of DIP. A treatment of these words as integrated loans would signify that DIP does contain retroflex phonemes, albeit restricted to words of non-Portuguese origin; if, on the other hand, these are interpreted as more or less sporadic loans, one would then conclude that retroflexes are entirely absent from the language's phonology. This is admittedly a difficult decision to make, and therefore my verdict that DIP does not accommodate retroflex phonemes - see Table (5.1) below - needs to be motivated. A number of facts conspire in support of this, including the following:

---

any language other than Portuguese, to the extent that one can say a South Asian etymological element is virtually absent. Almost all Gujarati or Indian English loans belong to open word classes.

    [3]It is of course possible that DIP *fuʈbɔl* derives from Portuguese *futebol* (without a retroflex) rather the Indian English etymon suggested here. If that were the case, however, this would be the only Portuguese-derived word in DIP to have a plosive segment reinterpreted as a retroflex; furthermore, the availability of Indian English to present-day native speakers of DIP (see section 2.2), which is responsible for many borrowings in their speech, turns Indian English into a very likely source.

a) all lexemes containing a retroflex sound are either Gujarati or English; with respect to the distribution of retroflex phonemes at least, there are then two separate lexical classes, one of which (that consisting of Ptg.-derived items) turns out to be dominant in DIP;

b) most lexemes containing a retroflex have a common Portuguese-derived correspondent in DIP (e.g. DIP *mɛdroz* for *bəɖka)*;

c) lexemes containing a retroflex are phonologically very similar to the items in the source language, i.e. there is no phonological adaptation to a DIP matrix;

d) certain cognates of lexemes with a retroflex occur in DIP without the retroflex; e.g. [fuʈ'bɔl][4] may also occur as [fut̚'bɔl], with an unreleased[5] alveolar plosive instead of a retroflex; one striking example of phonological adaptation, which simultaneously shows the aversion of DIP to both retroflexes and consonant geminates within a lexeme, is the common DIP word /da'ba/ 'box', from Gujarati /ɖəbbo/ 'box' (adapted from Mistry 1997:670).

The claim that the core phonological system of DIP does not accommodate retroflex sounds is particularly powerful in explaining patterns of phonological variation such as the one shown in d); because Indian English is available to many native speakers of DIP, the lexeme used to refer to 'football' may then be selected through one of two available media: a) the sporadic use of a word from Indian English (in which case no phonological adaptation occurs) or b) through resort to an item that has been borrowed into the lexicon of DIP (in which case phonological adaptation applies). This short discussion exemplifies some of the challenges in defining the phonological inventory of DIP, as well as the type of reasoning behind most of the decisions made (e.g. the absence of phonemic /h/).

The following sections define which segments have phonemic status in DIP and describe the most salient phonetic processes at play in shaping the spoken output as recorded in the corpus.

5.1 Vowel system

The vocalic phoneme system of DIP is organised in an oral and a nasal series with a certain degree of asymmetry among them, and they will therefore be considered separately. Nasality also intervenes in the definition of the language's diphthong inventory, so that we find both oral and nasal diphthongs, as described in 5.1.7.

The largest number of distinctive levels of height is attested for stressed oral vowels, which occupy a vocalic space with four levels (here defined as low, mid-low, mid-high and high) and make use of both the front, central and back positions. This four-way height system, however, is not relevant for all of the vocalic series, as will be seen below. DIP

---

[4]In accordance with the principles of the International Phonetic Alphabet, I use an apostrophe to signal primary stress in the phonetic transcription of words larger than monosyllables. The stress mark precedes the stressed syllable; see the *Handbook of the International Phonetic Association* (1999).

[5]See section 5.2.1 for an explanation of the unreleased nature of the plosive.

vowels are typically *peripheric*, i.e. front vowels are unrounded while back vowels are rounded.

Finally, the distribution and phonemic status of vowels is sensitive to considerations of lexical stress; distributional constraints operating on vocalic segments will be described in this section, but for a full account of DIP stress assignment refer to section 5.4 below.

5.1.1 Inventory of oral vowels

Figure 5.1 charts the DIP oral vowels with phonemic status, as demonstrated by the minimal pairs and further considerations below.
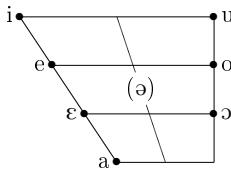
Figure 5.1: Phonemic inventory of DIP: oral vowels

The following minimal pair oppositions support most of the phonemic distinctions proposed in Figure 5.1 affecting segments contiguous in the chart:[6]

- - i vs. e: [dis] 'said'
                 [des] 'of this'

- - e vs. ɛ: [ped] 'fart'
                 [pɛd] 'stone'

- - ɛ vs. a: [ɛrf] 'grass'
                 [arf] 'tree'

- - a vs. ɔ: [bas] 'suffice'
                 [bɔs] 'dung'

- - ɔ vs. o: [nɔf] 'nine'
                 [nof] 'new'

No minimal pair of this type was found to illustrate the opposition between /o/ and /u/ but some close pairs can be produced in which these two segments occur in similar

---

[6]With respect to the phonetic representation of [ɛrf] and [arf], see section 5.2.4 on the final devoicing process affecting fricative segments.

phonological contexts, such as for example /pod/ 'rotten' vs. /tud/ 'all'. Furthermore, in stressed position at least, no allophonic variation ever obtains between the two.

The evidence for the phonemic status of /ə/, which stands for a cluster of allophonic mid-central realisations including the schwa and [ɐ], is less straightforward. One of the reasons for this is that /ə/ never occurs in stressed syllables, unless it is part of a diphthong.[7] Apparent minimal pairs can still be produced to justify its inclusion in the phonemic chart, although one must be weary of the fact that in all of the following what is being compared is a grammatical element with weak or no lexical stress (the prepositions *də*, *pə* and *nə*) and a stress-carrying lexeme:

- ə vs. a: [də] 'of'
          [da] 'give'

- ə vs. ɛ: [pə] 'to'
          [pɛ] 'foot/leg'

- ə vs. ɔ: [nə] 'in'
          [nɔ] 'knot'

The vocalic segment in these grammatical items has never been observed to occur in speech as anything but a schwa or a slightly lowered, allophonic [ɐ].

5.1.2 Stressed oral vowels

The lexemes that constitute the minimal pairs in the previous section, with the exception of those containing /ə/, exemplify the occurrence of all phonemes in stressed position. When stressed, none of these vowels allows significant allophonic variation.

There is however a phonetic process affecting mid-high and mid-low vowels in the speech of certain speakers, namely the insertion of a preceding glide (unless they are preceded by a phonological glide, in rising diphthongs, see 5.1.7). This glide takes the form of a palatal approximant before front vowels (e.g. [es] 'this' may be realised as [jes]) and a labial-velar approximant before back vowels (e.g. [a'ros] 'rice' may be pronounced [a'rwos]). This phenomenon was also noticed by Schuchardt's data collector, as some instances of <v> insertion occur in his 1883 description. Assuming that this refers to the same phenomenon just described for modern DIP, and although Schuchardt (1883) treats these segments as part of the lexemes, it is best to describe glide-insertion of this type as a phonetic- rather than phonemic-level operation, because a) it is strictly optional; b) it is more prevalent in

---

[7]The only exception is a grammatical item, the demonstrative *ikəl* 'that'. In its grammatical function (as a demonstrative as well as a marker of definiteness, see 7.2.6), this element carries weak or no stress and is subject to phonological bleaching, often being realised as [kəl]. However, the word maintains the /ə/ (pronounced schwa or [ɐ]) even if it stands as the sole constituent of an utterance, such as in the reply to a question of the type 'Which – –?'.

certain speakers than others; and c) it produces uncommon sequences such as glide-vowel-glide (e.g. [dwoj] for /doy/ 'two', or [wojʧ] for /oʤ/ 'today', see also 5.1.7).

5.1.3   Unstressed oral vowels

The realisation of DIP vowels is sensitive to lexical stress. Not only is the distribution of certain phonemes constrained by stress assignment; certain phonemic distinctions may also be suspended in unstressed positions. Absence of stress has two general consequences on DIP vowels, namely a) the annulment of the mid-high/mid-low contrast, and b) a heightening pull affecting certain segments. The particulars of these processes are described below.

One case in which the absence of stress dissolves phonemic boundaries concerns the series of back vowels, which was described in Figure 5.1 as consisting of /u/, /o/ and /ɔ/. In unstressed positions, however, phonetic [u], [o] and [ɔ] seem to be allophonic, but because of the overall dominance of [u] one may apprehend that absence of stress favours a phonetic shift towards the closed segment. The directionality of this shift is further supported by an important fact conditioning allophonic variation, namely that while certain lexemes allow free allophonic variation across the spectrum, others consistently employ [u]. To capture this crucial distinction, and considering that in unstressed position there are no grounds for a phonemic contrast between /o/ and /ɔ/, the spelling of lexemes containing unstressed back vowels with this type of variation in height (and which are often realised as [u]) will include either <o> or <ɔ> depending on its most common realisation. Examples of these two categories include:

a) allowing allophonic variation:
  - /bo'nit/ 'beautiful, good': [bunitˀ] / [bonitˀ] / [bɔnitˀ]
  - /tɔ'mo/ 'took': [tumo] / [tomo] / [tɔmo]

b) disallowing allophonic variation:
  - /fu'ʤi/ 'to escape': [fu'ʤi]
  - /u'ʋa/ 'to fly': [u'ʋa]

The situation concerning the front vowel series and its reaction to lack of stress is somewhat similar, but with some important differences. It is also the case that, in unstressed position, [e] and [ɛ] are in allophonic distribution (they are graphed <e> or <ɛ> in accordance with their most common realisation). The heightening pull, however, is weaker than that observed for the back vowels. Occurrences of this type of variation are indeed encountered in the corpus (e.g. /feʧad/ 'closed' may be realised as [fɛʧadˀ], [feʧadˀ] or [fiʧadˀ])[8] but so are numerous examples of lexical items that resist heightening (e.g. /pɛga/ 'to catch'). The phoneme [i], mirroring the case of [u], is consistently realised as /i/.

---

[8]Also [fɛʃadˀ], [feʃadˀ] or [fiʃadˀ]), see section 5.2.4 on the realisation of affricate segments.

The low vowel /a/ occurs in unstressed position and consistently takes its corresponding phonetic form [a]. It is therefore not affected by any heightening pull set in motion by the absence of stress. Examples include [la'rãjʃ] 'orange' and [na'se] 'to be born'.

It has already been pointed out that /ə/ may be realised as [ə] or [ɐ] and that it is limited to unstressed positions - but see this chapter fn. 7 -, be it within unstressed grammatical items (such as *də*, *pə* and *nə*) or in pretonic syllables of a stress-bearing lexeme. Examples of the latter include the words /də'pəj/ 'after(wards)', /pursə'sãw/ 'procession' or /kurə'sãw/ 'heart'. See section 5.4 below for an explanation of the prevalence (and deletion) of schwa in medial syllables of polysyllabic words, as in *kurə'sãw*.

5.1.4 Inventory of nasal vowels

Even though vocalic nasality is at times the by-product of specific phonemic environments (e.g. in the presence of a velar nasal, see section 5.2.2), the following minimal pairs - opposing homorganic oral and nasal vowels as well as oral and nasal diphthongs - make it clear that these are phonemically distinct in DIP:

- o vs. õ: [ot̚] 'other'
           [õt̚] 'yesterday'

- aj vs. ãj: [maj] 'more, but'
           [mãj] 'mother'

- i vs. ĩ: [si] 'if'
           [sĩ] 'yes'

We may still add the following non-minimal pairs in which, however, homorganic oral and nasal vowels occur in similar contexts:

- oj vs. õj: [foj] 'went'
            [põj] 'put(s)'

- u vs. ũ: [mi'nut̚] 'minute'
            [d͡ʒũt] 'together'

The nasal vowel series is reduced in comparison with the oral system presented in Figure 5.1 above. No distinction is made between mid-high and mid-low nasal vowels, which occur in free variation. In addition, no nasalised mid-central vowel occurs in DIP - cf. Figure 5.2.

On account of their relative rarity, it is considerably more difficult to find minimal pairs to illustrate the phonemic contrasts in Figure 5.2 than it was for the oral series, and it is
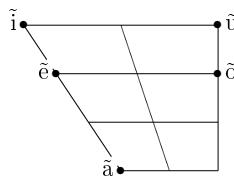
Figure 5.2: Phonemic inventory of DIP: nasal vowels

a matter of chance to be able to produce such pairs for contiguous segments on the articulatory chart. Not all of the following minimal pairs therefore reflect contiguity, but they are instrumental in defining the subset of segments in the nasal series with phonemic status:

- ĩ vs. ẽ: [sĩ] 'yes'
           [sẽ] 'without'

- ẽ vs. ã: [kẽt˺] 'hot'
           [kãt˺] 'sing(s)'

- ẽ vs. õ: [kẽt˺] 'hot'
           [kõt˺] 'tell(s)'

- ẽ vs. ũ: [ʤẽt˺] 'people'
           [ʤũt˺] 'together'

- ã vs. õ: [ãt˺] 'before'
           [õt˺] 'yesterday'

5.1.5 Stressed nasal vowels

All the examples provided in the previous section consist of stress-carrying monosyllables, which shows that all nasal segments identified in Figure 5.2 occur in stressed position. As noted before, the mid-high/mid-low contrast is not phonemic in the nasal series: the corresponding phoneme is variously realised as mid-high or mid-low, in free variation. As in the case of the oral series, stressed nasal vowels /ẽ/ and /õ/ may be preceded, in the speech of some speakers, by a glide. A palatal approximant may be inserted before a front vowel (e.g. /sẽp/ 'always' may be realised as [sjẽp˺]) and a labial-velar approximant may precede a back vowel (e.g. /õt/ 'yesterday' may be pronounced [wõt˺]).

5.1.6 Unstressed nasal vowels

All five nasal segments are admitted in unstressed position, as shown by the following forms:

[ĩ'vɛr] 'winter'
[sẽ'ta] 'to sit'
[jã'ta(r)] 'lunch'
[kõ'ta] 'to tell'
[ʤũ'ta] 'to gather'

In the absence of stress, though, nasal vowels may sometimes be denasalised in actual flowing speech, a phenomenon which does not usually affect their stressed counterparts. It is therefore not uncommon for /sĩ'ti/ 'to feel', for instance, to occur as [si'ti] or /kã'tig/ 'song' to be realised as [ka'tig̃]. Once denasalised, these vowels are subject to the heightening tendency typical of unstressed oral vowels (see 5.1.3), so that, for example, /ẽkə'tro/ 'met/found' may either be pronounced [ekə'tro] or [ikə'tro].

5.1.7 Diphthongs

In DIP, the vast majority of diphthongs occur in stressed position, although some counterexamples do exist (e.g. [kuj'dad̃] 'care' or [koj'tad̃] 'disgraced'). Both oral and nasal diphthongs are attested in DIP. The vast majority of the attested diphthongs are of the dropping type; the inventory of these, containing elements of both the oral and nasal series, is given below:

a) iw e.g. [viw] 'saw'
b) ew e.g. [na'sew] 'was born'
c) ej e.g. [ʋej] 'came'
d) ɛj e.g. [ʋɛj] 'old'
e) ɛw e.g. [sɛw] 'sky'
f) əj e.g. [də'pəj] 'after/then'
g) aj e.g. [ʋaj] '[to] go'
h) aw e.g. [paw] 'stick'
i) ɔj e.g. [mɔj] 'ground'
j) oj e.g. [kojs] 'thing(s)'
k) uj e.g. [mujt̃] 'very'
l) ẽj e.g. [ṽej] 'come(s)'
m) ãj e.g. [mãj] 'mother'
n) ãw e.g. [ʤam'lãw] 'a type of fruit'
o) õj e.g. [põj] 'put(s)'

It will be noticed that not all phonemes form diphthongs with both glides. The absence of */ij/ and */uw/ is articulatorily and perceptually justified, as the two segments in each diphthong are too close to each other. The absence of */əw/ should not be surprising -

/əj/ itself is rather limited in its occurrence. The real absence is that of */ow/, but see section 9.1.3 for a discussion of this issue.

Many instances of dropping diphthongs are the result of the fact that all etymological palatal lateral approximants have been reinterpreted in DIP as palatal glides. Examples include DIP [ʤu'ej] 'knee' (from Pt. *joelho* [ʒu'eʎu]) and DIP [ʊɛj] (from Pt. *velho* ['vɛʎu]).

Some attested rising diphthongs must be attributed to the operation of a phonetic principle which can be described as follows:

- insert a palatal glide between a back or central vowel and either a postalveolar fricative or a velar nasal.

The result of this process is predictable and stable. Examples include [lɔjʃ] 'shop', [lõjʃ] 'far', [ãjʃ] 'angel' and [ũjɰ] 'nail'. The insertion of a glide in these phonetic contexts serves the purpose of approximating the nucleus of the syllable to the point of articulation required for the coda consonant.

Rising diphthongs pose some important questions, as they may be the result of several phonetic processes. Two in particular are relevant in this case:

A - The optional *insertion* of a glide before stressed middle vowels (e.g. [kworp], [sjẽp]) - as described above.

B - The *diphthongation* of adjacent vowels forming the nucleus of two different syllables. Instances of verbal flexion (see section 6.2) exemplify this process: roots which end in a vowel (such as *pasi-* 'to travel' or *du-* 'to ache'), when combined with a vowel-initial suffix (e.g. the past suffixes -*o* or -*ew*, the infinitive suffixes -*a* or -*e*), often result in a rising diphthong; therefore, /pasi'a/ 'to travel' is realised as [pa'sja], /pasi'o/ 'travelled' surfaces as [pa'sjo], /du'e/ 'to ache' is realised as [dwe] and /du'ew/ 'ached' as [dwew].

These two processes operate in fundamentally different ways; notice that A concerns the *addition* of phonetic material absent from the phonological representation of the word whereas B entails the phonetic *alteration* of phonological material. However, they both produce glides that are strictly confined to phonetic realisations. The question then remains whether rising diphthongs are ever phonological.

The first person pronoun [jo] is a highly frequent instance of a rising diphthong for which there are no morphological means to support syllabification.[9] Other rising diphthongs are found in the corpus that contradict the predictions of process A, such as /si'kwẽt/ 'fifty', so these could be assigned to process B. However, velar stops are particularly common as precedents of rising diphthongs with /w-/, and this is also the case in Portuguese.[10]

---

[9]The transformation of the Ptg. etymon *eu* [ew] 'I' into a rising diphthong is exceptional, especially because the original dropping dipthong is present elsewhere in DIP (e.g. *dew* 'gave'). The unusual phonetic contour of DIP [jo] presumably arises from the abovementioned process by which a palatal glide is often inserted prior to a mid-front vowel; therefore a putative original form *ew came to be pronounced *[jew], after which the section [-ew] monophthongised in [o], resulting in [jo].

[10]For a similar debate concerning the phonemic status of Portuguese glides in rising diphthongs, see Mateus and d'Andrade (2000:48-51). These authors interpret all such diphthongs strictly as the product of phonetic principles, but they survey antagonistic opinions such as that of Barbosa (1965).

Therefore, for certain DIP lexemes a rising diphthong is etymological; examples include DIP [kwɔ'rẽtˀ] 'fourty' (from Pt. *quarenta* [kwɐ'rẽtɐ]), DIP [i'gwɔl] 'equal' (from Pt. *igual* [i'gwal]) and DIP [kwõtˀ] 'how many/much' (from Pt. *quanto* [kwẽ'tu]). Unfortunately, these cases do not contribute decisively to decide whether modern DIP produces such rising diphthongs on a phonemic level, as even these forms may sometimes be realised without the glide (e.g. [kɔ'rẽtˀ], [i'gɔl] and [kõtˀ]) - an optionality akin to that of the products of A for which no etymological glide obtained (e.g. DIP [kworpˀ] 'body', from Pt. *corpo* ['korpu]). It is clear, however, that DIP received these lexemes from Portuguese *with* the diphthong; notice that the etymological central vowels have been transformed into back vowels to accommodate the DIP requirement that /w/ may not precede central vowels. From a diachronic perspective, then, one can say that DIP rising diphthongs with a labial-velar glide overall spring from two distinct sources: a) the application of phonetic principle A, and b) the retention of etymological diphthongs in some instances preceded by a velar stop. It remains unclear whether this distinction is still relevant in modern DIP or whether all such diphthongs are now interpreted, by analogy, as the result of A.

5.2 Consonant system

Table 5.1 charts the consonantal segments with phonemic status in DIP according to their point and mode of articulation. The following sections analyse the distributional constraints and allophonic variation affecting consonantal phonemes, and in so doing motivate the contrasts proposed in Table 5.1. In the interest of clarity, the discussion is segmented according to mode of articulation.

	Bilabial	Lab-dent.	Dental	Alveolar	P-alveo.	Retrof.	Palatal	Velar	Uvul.	Glot.
Plosive	p　　b			t　　d				k　　g		
Nasal	m			n				ŋ		
Trill/Flap				r						
Fricative		f　　v		s　　z	ʃ					
Affricate				tʃ　　dʒ						
Approx.							j			
Lab. appr.								w		
Lateral				l						

Table 5.1: Phonemic inventory of DIP: consonants

5.2.1 Plosives

The following minimal pair oppositions motivate the proposed phonemic contrasts in the plosive series. The first observation to be made is that voice has phonemic value. Minimal pairs opposing two homorganic plosives distinguished solely by the voice parameter include the following:

- p vs. b: [patˀ] 'duck'

[bat˺] 'strike(s)'

- t vs. d: [tə] 'am/is/are'
  [də] 'of'

No exact pair has been identified for the velar pair in initial position, but consider the following pair in which they occur in an equivalent phonemic context:

- k vs. g: [kad˺] 'each'
  [gat˺] 'cat'

All plosives occur in syllable onsets word-initially, as demonstrated above, and also word-medially. Examples of the latter include [api'ŋa] 'to get', [ka'bo] 'finished', [ĩ'ter] 'whole', [a'diʋ] 'fox', [a'ki] 'here', and [a'gɔr] 'now'. They can also occur as syllable codas either word-medially or word-finally, but in these cases they are normally unreleased (the possibility exists for speakers to release such plosives audibly in formal elicitation contexts, but this hardly ever occurs in natural speech). An unreleased realisation neutralises the phonetic contrast between a voiced plosive and its homorganic voiceless counterpart, but not their phonological opposition, as demonstrated by the pair /jak/ 'jackfruit' vs. /jag/ 'jaggery'. In concatenated speech the final consonant will be fully released if the following segment is vocalic and the voice contrast therefore re-surfaces. Compare the behaviour of the velar plosives in the following examples:

(15) a. aki fik - [aki'fik˺]
      b. fik aki - ['fika'ki]
          'stay here'

(16) a. a el dig - [a'jel'dig˺]
      b. dig a el - ['diga'jel]
          'tell him'

5.2.2 Nasals

Minimal pairs can be produced to attest of the phonemic distinction between the bilabial and the alveolar nasals:

- m vs. n: [mãw] 'hand'
  [nãw] 'no'

One might want to interpret the occurrence of nasal vowels as the result of spread of nasality from a phonemic nasal consonant, as proposed for several languages. This is not

the case in DIP; the presence of a nasal consonant in the coda does not necessarily entail the nasalisation of the previous vowels, as conveniently demonstrated by the following examples:

- ĩ vs. im vs. in: [sĩ] 'yes'
                    [i'sim] 'on top'
                    [fin] 'small'

This is not to say, however, that nasality spread cannot occur. In fact, that is often the case when the following nasal consonant is /ŋ/. The phonemic status of this segment is somewhat complex; its production in certain contexts must be seen as belonging to the realm of articulation alone, viz. as a linking phone between a nasal vowel and a velar plosive (e.g. [brãŋk˺] for /brãk/ 'white'). On the other hand, /ŋ/ occurs distinctively in phonological contexts in which another nasal consonant or a single nasal vowel are licensed; compare, for instance, [sĩ], [i'sim] and [fin] with [tiŋ] 'had/there was' or the use of the suffix /-iŋ/ (e.g. [ra'tiŋ] 'small mouse', [au'ziŋ] 'grandmother' - see section 8.2.1). Articulatorily speaking, /ŋ/ has two allophones: it may surface either as a velar nasal, before consonants (e.g. [tiŋku'me] for /tiŋ ku'me/ 'was eating'), or as a velar approximant /ɰ/ intervocalically (e.g. [ki'ɰẽt] 'five hundred', [raj'ɰo] 'scratched') and in isolation (e.g. [tiɰ] for /tiŋ/ 'had/there was'). An important distinction between phonetic and phonemic ŋ is that the latter, as described in 5.1.7, triggers the realisation of a palatal glide before it when preceded by a vowel other than /i/ (the sequence [raj'ɰo] above therefore corresponds to /raŋo/ on the phonemic level).

5.2.3   Trill and flap

Table 5.1 makes no distinction between an alveolar trill and a flap, although both occur in actual speech. A closer look at their distribution suggests that they are allophones. The flap occurs as the second element of consonant clusters (e.g. [pɾet] 'black'), in syllable codas (e.g. [gaɾ'gãt] 'throat') and syllable onsets in word-medial positions (e.g. [ma'ɾad] 'built'). The flap is dispreferred in word-initial positions, and the trill is blocked from syllable codas. In addition, speakers show free variation between the flap and the trill in word-medial syllable onsets, with a clear preference for the flap in a token count.

All indications taken together, the trill and the flap are here interpreted as allophones of an alveolar phoneme /r/. No minimal pairs can be found in which /r/ appears as the second element of a consonant cluster for the simple reason that this is one of the few phonemes allowed in such a position. The phonological nature of /r/ is made evident from the following examples in which the segment appears word-initially, word-medially and word-finally:

- r vs. p: [rɔd˺] 'wheel'
               [pɔd˺] 'can v.'

- r vs. m: [ku're] 'to run'
               [ku'me] 'to eat'

- r vs. z vs. d: [kar] 'expensive'
               [kaz] 'house'
               [kad] 'each'

5.2.4   Fricatives and affricates

Regarding the labiodental point of articulation, the true opposition on the phonetic level is not between a voiced and voiceless fricative but rather a voiceless fricative and a voiced approximant.

- f vs. ʋ: [fak̚] 'knife'
             [ʋak̚] 'cow'

Even though full labiodental constriction is never attained, the latter segment will be treated as the phonetic realisation of a fricative /v/ in order to

a) distance it from the series of approximants - see below (this section) for a discussion of the contrast between /v/ and /w/;

and

b) highlight its similarities with the fricative series, such as the fact that /v/ is also subject to final devoicing, in which case it is realised as a voiceless labiodental fricative [f].

The approximant realisation of this fricative is at odds with DIP's main lexifier, but it is easily explained by looking at the phonetic and phonemic systems of its surrounding languages. The labiodental approximant (but not the voiced fricative) is in fact a common segment in Western Indo-Aryan languages, and DIP sides with them in this respect. However, in these languages the approximant [ʋ] and the labial-velar glide [w] are in allophonic distribution, with the former being reserved as precedent to a front vowel (see Masica 1991:100). In DIP these two segments do contrast, although the evidence is limited. One of the reasons why it is impossible to obtain exact minimal pairs has to do with distributional constraints affecting these two segments: whereas [w] never occurs at syllable onset except as the second element of complex onsets, [ʋ] only rarely occurs immediately after a vowel (although it may occur in codas, consider /arv/ 'tree', for instance). Among the few cases in which it does, the following pair was selected to illustrate that [ʋ][11] and [w]

---

[11]It will be made clear at a later stage that in syllable-final position, voiced fricatives are unspecified as to their voice trait, and are only realised as voiced if the segment that follows is voiced. In isolation, this lexeme would be pronounced [laf]; the phonetic representations given in this section are necessary for the argument and therefore assume non-isolated occurrences of the words in question.

occur in similar contexts:

- ʋ vs.w: [laʋ] 'wash(es)'
             [paw] 'stick'

Some additional instances of postvocalic /ʋ/ include [aʋˈziŋ] 'grandmother', [oʋ] 'egg' and [nɔʋ] 'nine'.

Section 5.1.7 above describes how postalveolar fricatives trigger the insertion of a glide [j] after a preceding non-front vowel, giving rise to a number of phonetic diphthongs. Postalveolar fricatives, both voiced [ʒ] and voiceless [ʃ], are phonetically prominent in the corpus, but a closer look at patterns of variation reveals that there is more than meets the eye. In most of the cases they are allophones of their affricate counterparts [ʤ] and [ʧ], as both realisations are licensed. Consider the following pattern of phonetic variation: [ʃuʋ] vs. [ʧuʋ] 'rain', [ʒũt˺] vs. [ʤũt˺] 'together'. Both are also possible if the segment occurs in syllable coda, although in this position the postalveolar fricative is preferred; examples include the variable forms [lɔjʤ] vs. [lɔjʒ] 'shop'. Such preference is understandable - indeed expected - in light of the Sonority Sequence Principle (v. Selkirk 1984). This phonotactic principle predicts that the syllable nucleus constitutes a sonority peak to which the onset progressively builds up and from which the coda progressively descends; plosives are by definition the least sonorant in the scale, and as such the neutralisation of the obstruence in the syllable-final affricate ensures that the final segment is the one least sonorant in the coda.

A small number of lexemes always surface with [ʃ] and disallow [ʧ], most notably forms of the highly frequent verb /diˈʃa/ 'to let'. As such, a separate phoneme /ʃ/ has to be recognised, distinct from /ʧ/ on the phonological level but not necessarily so on the phonemic level. Interestingly, this opposition has no parallel for the corresponding voiced segments, as [ʒ] and [ʤ] are strictly allophones; this creates an instance of asymmetry as far as postalveolar fricatives are concerned.

We have now established that phonetic [ʃ] may correspond to phonemic /ʃ/ or /ʧ/, but in a limited subset of the lexicon it can still be at variance with the alveolar fricative [s]. These include occurrences of the fricative in syllable-final word-internal position, producing the following pairs of phonetic realisations: [raʃˈpa] vs. [rasˈpa] 'to scratch'; [iʃˈkɔl] vs. [isˈkɔl] 'school'. The postalveolar realisation of sibilants in syllable-final position is a distinctive trait of Modern European Portuguese, but not of the 16th-century varieties (see section 9.1.3), and must therefore be a relatively recent trend in DIP.

A very clear tendency towards final-devoicing affects DIP fricatives in a particularly strong way, to such an extent that the phonemic opposition between voiced and voiceless fricatives is suspended in this position. In isolation, all words ending in a fricative surface with a voiceless consonant; if the underlying fricative is voiced, however, they will become voiced if the subsequent segment is also voiced. The following examples concern the adjective /noʋ/ 'new'; they illustrate how the final voiced phoneme is voiceless in isolation (17a) and before another voiceless phoneme (17b) but voiced if the following segment is also voiced, whether it is homorganic (17c) or not (17d):

(17)   a.   /nov/ 'new'- [nof]

        b.   /nov kaz/ 'new house'- ['nof'kas]

        c.   /nov vistid/ 'new dress'- ['noʋʋiʃ'tid]

        d.   /nov irmãw/ 'new brother'- ['noʋiɾ'mãw]

Final devoicing similarly affects the remaining fricatives, so that under similar circumstances [kas] 'house' occurs instead of [kaz] and [lɔjʃ] 'shop' instead of [lɔjʒ]. This pattern is not without exceptions, as words may consciously be pronounced with a voiced final fricative in contexts where one might expect devoicing. The truth remains, however, that the general pattern is the one illustrated by the contextual variation of /vez/ in the following utterance by a young male speaker (phonetic transcription):

(18) [*jo tə   kai ãŋu ʋez   ãŋu ʋes*]
      1s HAB fall sometimes sometimes
      'I fall down sometimes, sometimes.'

5.2.5   Approximants

It is convenient to introduce the section on approximants with a reminder that some non-approximant phonemes routinely surface as approximants. The most common realisation of the velar nasal /ŋ/ is a velar approximant [ɰ] but, as argued in 5.2.2, there is no phonemic opposition between the two. The same holds for the labiodental approximant [ʋ] which, as explained in section 5.2.4, is best treated as representing an underlying fricative /v/.

With these phones excluded from the phonemic chart, it can now be established that DIP has two approximant phonemes: palatal /j/ and labial-velar /w/. They always combine with vowels to form diphthongs, as described in section 5.1.7 above. Examples of minimal pairs opposing these glides to each other and to other consonants are given below:

  - j vs. w: [paj] 'father'
              [paw] 'stick'

  - j vs. d: [mej] 'half'
              [med̚] 'fear'

  - w vs. g: [diw] 'Diu'
              [dig̚] 'say(s)'

Audible glides occur as the result of several phonetic processes and are therefore not instances of the corresponding phonemes. A phonetic [w] can be produced before a stressed

middle back vowel (e.g. [wõtˀ] for /õt/ 'yesterday') or as the realisation of unstressed /u/ when followed by another vowel (e.g. [dwe] for /du'e/ 'to hurt'). Similarly, a non-phonemic [j] may also precede a stressed middle front vowel (e.g. [pjɛdˀ] for /pɛd/ 'stone') or be the realisation of a diphthongised /i/ followed by another vowel (e.g. [lji̯ãw] for /li'ãw/ 'lion'); in addition, a palatal approximant may mediate between most vowels and the postalveolar fricative [ʃ] or the velar nasal [ŋ]. For more examples and further information, see section 5.1.7.

5.2.6   Laterals

The phonological repertoire of DIP features one lateral phoneme whose articulation is alveolar except in word-final position, where it tends towards a dental realisation. The minimal pairs presented below attest to its phonemic status, while making it clear that /l/ is licensed in both syllable onsets and codas:

- l vs. s: [mil] 'a thousand'
                [mis] 'mass'

- l vs. p vs. k: [ladˀ] 'side'
                [padˀ] 'priest'
                [kadˀ] 'each'

It can also occur as the second element in a consonant cluster, such as in the word /flor/ 'flower'. In that respect, the distribution of /l/ approaches that of /r/ (cf. e.g. /fri/ 'cold'), but the following minimal pair clarifies that a phonemic contrast obtains between the two segments:

- l vs. r: [kal] 'shut up'
              [kar] 'expensive'

5.3   The syllable

DIP words may be monosyllabic, disyllabic or trisyllabic. Even though suffixation often involves the addition of a syllable, the resulting form never contains more than three syllables. The syllables admit simple and complex onsets and codas, whereas the nucleus may be occupied by up to two segments (see 5.3.2 for a full account). The following subsections describe in more detail the distribution of the various phonemes within the syllable.

One observation that must be made at this point concerns geminates that may possibly arise in discourse. Whenever similar segments (whether vocalic or consonantal) occur contiguously, DIP does not contract the resulting geminates. The phrase /kõ'pra a'roz/

'to buy rice' is pronounced [kõ'praa'ros] and /ain nə foj/ 'didn't go yet' as [a'innə'foj].

5.3.1 Onset

DIP admits simple and complex onsets. Below is a list of monosyllabic words with simple onsets:

/pork/	'pig'	/bat/	'hit(s)'
/tud/	'all'	/da/	'give(s)'
/kor/	'colour'	/gat/	'cat'
/mujt/	'very'		
/nad/	'nothing'		
/ru/	'street'		
/fin/	'small'	/vak/	'cow'
/sol/	'sun'	/zɛl/	'[a name]'
/lad/	'side'		
/ʧuv/	'rain'	/ʤa/	'already'

This list makes it clear that all consonantal phonemes are allowed in this position, with the exception of the velar nasal /ŋ/ and the glides. Whereas /w/ never occurs syllable-initially, there is one highly common word in which /j/ occupies the coda position, namely the 1st person singular personal pronoun /jo/ 'I'. See section 5.1.7 for a discussion of this exceptional case.

Complex onsets are occupied by consonant clusters, the second element of which must be a liquid; /r/ is preferably flapped in this position, and it is much more common than /l/. Onset clusters have either a plosive or a fricative as their initial element. Some attested cases are given below:

/pret/	'black'	/brãk/	'white'
/trej/	'three'	/drumi/	'to sleep'
/kruz/	'cross'	/graj/	'crow'
/fri/	'cold'		
/flor/	'flower'		

5.3.2 Nucleus

Section 5.1.7 claimed that the status of rising diphthongs (with the exception of /jo/) in cases such as [si'kwẽt] 'fifty' or [i'gwal] 'equal' is unclear. If they are the result of the diphthongation of adjacent vocalic segments, then such segments must be assigned to different syllables; if, on the other hand, they are phonological, then the rising diphthong integrates the syllable nucleus.

98

One might feel tempted to assign glides of rising diphthongs to the syllable onset (as elements in a consonant cluster); likewise, the glide in a dropping diphthong may be interpreted as a syllable coda. In this interpretation, the sequence /jz/ in /kojz/ 'thing' would be seen as a cluster. In the absence of clues to resolve this interpretative standstill, the issue will have to remain unresolved. As such, it is unclear whether the syllable nucleus can only contain a single segment or up to two segments.

5.3.3 Coda

Syllable codas in DIP can also be simple or complex. I assume, for the purpose of this description, that glides are assigned to syllable nucleus (see section 5.3.2 for a full discussion) and not to syllable coda.

All consonantal phonemes are allowed as simple codas but, as mentioned earlier, in this position there is no phonetic opposition between /f/ and /v/, and between /ʧ/ and /ʤ/. The contrast between /s/ and /z/ itself is not always kept in coda position, although certain lexemes (e.g. /vez/ 'occasion' or /purtə'gez/ 'Portuguese') do have a tendency to contradict the instructions of final devoicing.

/ap/ 'flatbread'   /kɔb/ 'snake'
/ot/ 'other'   /bud/ 'stupid'
/pok/ 'few'   /fig/ 'fig'
/tɔm/ 'take(s)'
/piken/ 'small'
/uŋ/ 'nail'
/mojr/ 'muslim'
/nɔv/ 'nine'
/dos/ 'sweet'   /vez/ 'turn'
/ʤanɛl/ 'window'
/oʤ/ 'today'   /pɛʧ/ 'flat'
/peʃ/ 'fish'

The consonant clusters allowed as complex codas are nearly a mirror image of those attested for complex onsets. In this case, the first element of the cluster (the most sonorant) must be a liquid and the final segment either a plosive or a fricative. There is at least one case in which the final consonant is nasal, viz. *karn* 'meat', although it only surfaces in the speech of some (cf. this chapter fn. 14). Once more, /r/ is normally flapped and it is more common than /l/ in the first position. Below are some attested examples of complex codas:

/korp/ 'body'
/pɔrt/ 'door'
/alt/ 'tall'
/pork/ 'pig'
/fors/ 'speed'

/bols/ 'pocket'
/arv/ 'tree'

5.4 Stress and intonation

DIP is a stress-accent language, in that each word marks one syllable as being particularly prominent (see e.g. Bolinger 1958, Smith 2008b). Stress falls predictably on the word's final syllable, and as such can be said to be phonologically assigned rather than morphologically determined. This claim is based on a number of different considerations:

- *Phonetic*: in polysyllabic words, final syllables are consistently more prominent than others in terms of pitch.[12] Intensity also clearly favours a given word's last syllable most of the times; whenever it does not, the non-final syllables seldom supersede it, in any context;

- *Phonological*: as explained in section 5.1, the phonemic contrast between /e/ and /ɛ/, and between /o/ and /ɔ/ only applies to final syllables. The distribution of /ə/ is also sensitive to stress in that it is excluded from stressed syllables (with the exception of *ikəl* 'that', *dəpəy* 'after(wards)' and certain English-derived words such as *saykəl* 'bicycle', see above);

- *Morphological*: the addition of morphological material onto verbal stems can increase the number of syllables in a word and is therefore a convenient testing ground for the claim that DIP stress is phonologically determined; as predicted, in such cases we do observe a fluctuation of stress from the last syllable of the stem onto the verbal suffix, as exemplified by the following pairs:[13]

- monosyllabic: /fal/ 'speak(s)'
  disyllabic: /fa.'lo/ 'spoke'

- disyllabic: /a.'paɥ/ 'get(s)'
  trisyllabic: /a.pə.'ŋo/ 'got'

- *Etymological*: in the adaptation of Portuguese words into DIP, etymological post-stressed syllables are dropped (with the potential incorporation of their onset as the coda of the last DIP syllable), with the effect that in DIP all final syllables coincide with stressed syllables in Portuguese.[14] Consider the following examples:

---

[12]This is a strong general pattern, but it can of course be subverted by certain prosodic contours and by sandhi rules dictating the spread of high pitch from a preceding word onto the word's first syllable.

[13]Whenever relevant, syllable boundaries are indicated by single dots in phonemic and phonetic representation. See also section 4.3.

[14]Very rarely, some speakers will add phonetic material to represent the last syllable of the etymon in such cases, so that 'music' may be pronounced [muz], [musk̚] or [muzk̚]; the same holds, for instance, for the Ptg. name 'Fátima', which is normally [fat̚] but may sometimes occur as [fatm]. While, undoubtedly,

- one post-stress syllable:	Ptg. /ku.'bɛr.tɐ/ 'blanket'
	DIP /ku.'bɛrt/ 'blanket'

- two post-stress syllables:	Ptg. /'mu.zi.kɐ/ 'music'
	DIP /muz/ 'music'

In disyllabic words, no significant vowel reduction is observed in the unstressed syllables, but the situation is rather different for trisyllabic words. In these cases, the first syllable shows little reduction when compared to the stressed syllable, but the second (medial) syllable typically undergoes considerable reduction. The nucleus of such syllables is usually occupied by /ə/, which will surface as [ə], [ɐ] and often [Ø].[15] Consider the following examples:[16]

- /kazə'mẽt/ 'wedding':   [kaz'mẽtˀ]

- /rapə'rig/ 'girl':   [ra'prigˀ]

- /kurə'sãw/ 'heart':   [kuɾ'sãw]

As the previous examples illustrate, reduction affects the vocalic segment of the medial syllable but not the remaining material, which is redistributed onto the surrounding audible syllables. This poses an important constraint on the tendency to reduce medial syllables, namely that the resulting consonant clusters should not violate the language's cluster restrictions. As shown in the following example, consonant clusters containing three segments are not necessarily blocked, but only occur in certain contexts:

- /purkə'ri/ 'mess':   [pur'kri] or [purk'ri]

Notice that two different syllable divisions are proposed for the phonetic realisation of this word, in turn assigning the velar plosive to the first syllable or the final one. Whatever the interpretation, the resulting clusters - onset plosive + liquid, or coda liquid + plosive - are found elsewhere in DIP (see section 5.3). The distribution of a sequence containing

---

such extra material must be generated at the phonological level in these examples, the speakers who produce the longer forms tend to be the ones most familiar with standard Portuguese; as such, it seems warranted to propose that such realisations represent a type of correction modelled on SP.

[15] Considering that stress always falls on the last syllable, it is striking that the first and second syllables of trisyllabic words behave so differently with respect to this tendency of vowel reduction. The reason for resistance of the first syllable to reduction may be psycholinguistic; it has been shown that the first and final syllables of a word are particularly important both for processing by a hearer and retrieving by a speaker - the so-called *bathtub effect* (see e.g. Scovel 1998) by analogy with a person lying in a tub with only her feet and head visible outside the water.

[16] The phonetic representation of [kuɾ'sãw] mirrors the record of the same word in a song as transcribed in Schuchardt (1883), in the verse *Regai mé corção* 'water my heart'; the orthography clearly represents a similarly ellipted medial syllable, which reveals that this type of reduction is not a recent development in DIP.

a liquid + plosive + plosive might be more difficult to resolve, which can explain why the middle syllable in /purtəˈgez/ 'Portuguese' is seldom reduced.

As far as intonation is concerned, it is clear that its assignment responds to grammatical and pragmatic considerations. In general, each clause has an intonational peak which is tendentially assigned to a focal constituent - which, if not the verb form, is in many cases expressed in preverbal position in declarative sentences (see sections 7.8.1 and 7.8.4). After this intonational peak, there is often an abrupt drop in pitch. There is therefore a close interaction between syntactic realisation and intonational contours which favours the preverbal position as a focus position and, therefore, a common locus for the clause's intonational peak. This need not be the case, however; in cleft constructions (see 7.8.1), for instance, the focused element and the VP may be mediated by other constituents, and it is the clause-initial clefted element which is prosodically most prominent.

Some grammatical elements (e.g. simple prepositions, many auxiliaries), typically unstressed or rather weakly so, do not receive high prosodic prominence; others (e.g. negators, complex prepositions) accept it, when they are focused (e.g. the negator in *nã ɛ* 'it is NOT', in contrast to the polarly opposed assertion *ɛ* 'it IS').

In content questions, the intonational contour is usually as described above: the pragmatically central element, i.e. the interrogative pro-form, is consistently realised in preverbal position, and it also carries the intonational peak (see section 7.6.2.3). The syntax of polar questions, on the other hand, is indistinguishable from that of simple declaratives (see 7.6.2.1), so they are distinguished strictly by intonation. The prosody of typical polar questions involves rising intonation, with its peak at the end of the clause. Tag questions (a particular type of polar questions, described in 7.6.2.2) behave differently in that their interrogative status is marked by the use of a tag *nə* which, in prosodic terms, is normally the most prominent element of the clause.

5.5 Orthographic conventions

The description of DIP phonology in this chapter now allows for the definition of orthographic rules that take phonemic contrasts in consideration - this writing system is based on the Roman alphabet for practical purposes. In so doing I shall have four main preoccupations:

a) the orthography should be consistent and mirror phonemic structure as closely as possible;

b) the orthography should facilitate the representation of orality; therefore, whenever phonetic variation occurs, the orthography adopted will correspond to the most common realisation;

c) the symbols should be easy to hand-write and type; therefore, resort to IPA symbols will be reduced to a minimum, although it is unavoidable to code certain phonemic oppositions;

d) the symbols should be easily apprehended so that the orthography may potentially be adopted by the speakers of DIP; in this sense, the proposed orthography tendentially relies on common transcription rules in place for the romanisation of Indo-Aryan languages, as well as on the official orthography of standard European Portuguese;

Below is an explanation of the orthographic rules adhered to in this volume, which amount to a writing system for DIP. It will be noticed that, unlike in the orthography of European Portuguese, the present system does not resort to diacritics, with the exception of the tilde to indicate nasality. This decision is made in order to avoid confusion, as graphic accents in SP not only clarify vowel quality (e.g. <ê> for [e] vs. <é>/<è> for [ɛ]) but often also indicate the position of word stress. In contrast, no visual indication of stress is adopted for DIP because, as described in 5.4, stress is phonologically determined and not morphological.

The following symbols stand for the oral vowels:

- <a> corresponds to /a/: e.g. *ap* 'flatbread'.

- <ɛ> corresponds to /ɛ/: e.g. *pɛd* 'stone'; this phoneme contrasts with /e/ in stressed syllables only, in unstressed syllables <ɛ> will be preferred over <e> to represent /e/ whenever the segment surfaces as [ɛ] more often than [e][17], such as in *sɛtẽt* 'seventy'; the compound <ai> usually employed in the romanisation of modern Indo-Aryan languages is not suitable because DIP does allow for the juxtaposition of /a/ and /i/, such as in *ain* 'still'.

- <e> corresponds to /e/: e.g. *es* 'this'; this phoneme contrasts with /ɛ/ in stressed syllables only, in unstressed syllables <e> will be preferred over <ɛ> to represent /e/ whenever the segment surfaces as [e] more often than [ɛ].

- <i> corresponds to /i/: e.g. *isim* 'on top'.

- <ɔ> corresponds to /ɔ/: *ɔm* 'man'; this phoneme contrasts with /o/ in stressed syllables alone, in unstressed syllables <ɔ> will be preferred over <o> to represent /o/ whenever the segment surfaces as [ɔ] more often than [o], such as in *kwɔrẽt* 'fourty'; the compound <au> usually employed in the romanisation of modern Indo-Aryan languages is not ideal as it might be misinterpreted as a diphthong after the Portuguese orthography.

- <o> corresponds to /o/: e.g. *ov* 'egg'; this phoneme contrasts with /ɔ/ in stressed syllables only, in unstressed syllables <o> will be preferred over <ɔ> to represent /o/ whenever the segment surfaces as [o] more often than [ɔ].

- <u> corresponds to /u/: e.g. *uva* 'to fly'.

---

[17] This type of reasoning may not always be straightforward, i.e. the corpus may not be large enough to decide on which of the realisations is the most common; in such cases I will prefer <e> and <o> over <ɛ> and <ɔ> respectively because the former are easier to reproduce by hand or on most keyboards.

- <ə> corresponds to /ə/, which can be pronounced [ə] or [ɐ]: e.g. *nə* 'in'; neither standard Portuguese or the romanisation scheme of Indo-Aryan languages contain a dedicated symbol for the schwa.

Nasal vowels are indicated by a tilde. The correspondence between a grapheme and a phoneme is very straightforward in this series. Notice that the mid-high/mid-low opposition does not apply for the nasal series, and the orthographic system reflects that reality:

- <ã> corresponds to /ã/: e.g. *amiŋã* 'tomorrow'.

- <ẽ> corresponds to /ẽ/: e.g. *dẽt* 'tooth/inside'.

- <ĩ> corresponds to /ĩ/: e.g. *ĩter* 'whole'.

- <õ> corresponds to /õ/: e.g. *õt* 'yesterday'.

- <ũ> corresponds to /ũ/: e.g. *jũt* 'with/together'.

This is the chosen notation for the semi-vowels:

- <y> corresponds to /j/: e.g. *gray* 'crow'; the IPA symbol is not selected here so as to reserve <j> for the voiced affricate (see below).

- <w> corresponds to /w/: e.g. *diw* 'Diu'; notice that the proposed orthography refrains from representing glides that result from phonetic processes.

The consonants of DIP are graphed as follows. One compound grapheme <ch> is proposed; the combination of a consonant with <h> is unequivocal in DIP because no aspiration ever occurs.

- <p> corresponds to /p/: e.g. *pimer* 'first, earlier'; the unreleased occurrence of plosives is purely phonetic and therefore not coded in the orthography.

- <b> corresponds to /b/: e.g. *buf* 'buffalo'.

- <t> corresponds to /t/: e.g. *tãt* 'so/as'.

- <d> corresponds to /d/: e.g. *doy* 'two'.

- <k> corresponds to /k/: e.g. *kabrit* 'goat'; this notation is preferred to <c> (the most common representation of /k/ in standard Portuguese) in that it avoids the peculiarities associated with SP <c>, which stands for /s/ before <e> or <i>; in addition, this correspondence is common in the romanisation of Indo-Aryan languages.

- <g> corresponds to /g/: e.g. *gran* 'big'; unlike in SP and English, <g> always stands

for /g/ in DIP, even when it occurs before <e> and <i>.

- <m> corresponds to /m/: e.g. *moyr* 'muslim'.

- <n> corresponds to /n/: e.g. *nad* 'nothing'.

- <ŋ> corresponds to /ŋ/: e.g. *apiŋa* 'find, get'; it is difficult to replace this IPA symbol with any other grapheme or sequence: the sequence <nh> might be interpreted as referring to a palatal rather than velar nasal (as in SP), and the sequence <ng> could potentially trigger wrong syllable break-ups.

- <r> corresponds to /r/: e.g. *ratiŋ* 'mouse'.

- <f> corresponds to /f/: e.g. *fɛs* 'party'; it does not occur in word-final position, where its voiced counterpart is employed (see 5.2.4 above for the neutralisation of the /f/ vs. /v/ opposition in such contexts).

- <v> corresponds to /v/, which consistently surfaces as [ʋ]: e.g. *vay* 'go(es)'; it is preferred to <f> in word-final position.

- <s> corresponds to /s/: e.g. *sapat* 'shoe'; it may, in syllable codas, be pronounced [ʃ] in allophonic variation with [s], as described in 5.2.4 above.

- <z> corresponds to /z/: e.g. *zasey* 'sixteen'; in word-final position, it is subject to devoicing in most contexts.

- <x> corresponds to /ʃ/: e.g. *dex* 'let'; <x> contrasts with <ch> (/tʃ/) in that the former is always pronounced [ʃ].

- <ch> corresponds to /tʃ/, which can be realised as [tʃ] as well as [ʃ]: e.g. *churis* 'sausage'; see section 5.2.4 for information on variability constraints; this notation is retrieved from SP and must not be mistaken for the aspirated affricate as the norms of romanisation of Indo-Aryan languages might suggest.

- <j> corresponds to /dʒ/, which can be realised as [dʒ] or [ʒ]: e.g. *janɛl* 'window'; see section 5.2.4.

- <l> corresponds to /l/: e.g. *lãbe* 'to lick'.

In order to illustrate the application of orthography as well as the phonemic-phonetic correspondences so far debated in this chapter, a short text is given below in phonetic, phonemic and orthographic form. This story was told by a male speaker of DIP who was 9 years old at the time of the recording (March 2005).

The phonetic transcription contains certain instances of hesitation (e.g. [i] or [ũ] in the sequence [i'dʒẽgli ũ'dʒẽgli'dʒɛt]) and a promptly corrected sequence of liquid inver-

105

sions (['latkil'iɲsaɾʋa] instead of ['ratkir'iɲsalʋa]) which are regularised in the phonological and orthographic representations. The text also contains borrowings from Indian English (namely *jungle*, *lion* and *jungly [jat]*), easily recognised for their initial stress pattern that contradicts the general rule described for DIP (see section 5.4); these lexemes are replaced by double dashes (– –) in the phonological transcription and italicised in the orthographic version.

a) phonetic transcription

[ũ'diũˈɾejdə'ʤẽglˈtiŋ | jũpi'ken'rat:iŋːˈʋajnɐdelkaˈmiŋ || iːˈɾejdə'ʤẽglfaˈlo | u'seɛˈtãtpi'ken | u'sewkiəfa'ze || i'ʤẽgli ũ'ʤẽgli'ʤɛt'ʋej | am'ro'pɛdikəlːajɔn || ɛrpəku'mea'jel || iːˈrat 'latkil'iɲsaɾʋa sal'ʋaːjel || i rat'fwoj | mur'dewmur'dewpikɐ'liŋ | sul'to | idə'pəj 'lajɔnfa'lo kiːˈmujtˈubriˈgadˀ|| iʃˈtɔrkaˈbo]

b) phonological transcription

/ũ di ũ rej də – – tiŋ | i ũ pi'ken rat tiŋ vaj nə del ka'miŋ || i rej də – – fa'lo | u'se ɛ tãt pi'ken | u'se uki a fa'ze || i ũ – – vej: | amə'ro pɛ di'kəl – – || ɛr pə ku'me a el | i rat ki'riŋ sal'va a el || i rat foj | mur'dew mur'dew pikə'liŋ | sul'to || i də'pəj – – fa'lo ki mujt ubri'gad || is'tɔr ka'bo/

c) orthographic transcription

<ũ di ũ rey də *jungle* tiŋ, i ũ piken rat tiŋ vay nə del kamiŋ. i rey də *jungle* falo "use ɛ tãt piken, use uki a faze?". i ũ *jungly jat* vey i aməro pɛ dikəl *lion*. ɛr pə kume a el. i rat kiriŋ salva a el. i rat foy, murdew murdew pikəliŋ, sulto. i dəpəy *lion* falo ki "mujt ubrigad". istɔr kabo.>

d) free translation

'Once upon a time there was a king of the jungle, and a little mouse was going about his way. And the king of the jungle told him "you are so small, what are you going to do?". And then a wild man came and bound the lion's leg, he intended to eat him. And the mouse wanted to save him. So the mouse went and gnawed and gnawed [at the rope] and set him free. And then the lion said "thank you very much". The story is finished.'

Chapter 6

Parts of Speech

The purpose of this chapter is two-fold. The first of these is to set out and motivate Parts-of-Speech (henceforth PoS), or lexical categories, onto which to assign the lexemes of DIP, therefore providing a descriptive array with which to structure the syntactic analyses in chapter 7. In this endeavour, I am guided as much by theoretical and crosslinguistic traditions as by a need for descriptive clarity and exhaustiveness. A full motivation of the criteria involved in defining word classes for DIP is given in section 6.1 below.

The second purpose of the present chapter is to describe the morphological operations available to DIP. Given that morphological considerations are often key to determining categorical assignment, it makes sense to discuss morphology on a category-by-category basis. A further reason for deciding to engage in a description of morphology here is that, when compared to its ancestor languages, DIP makes remarkably little use of morphological means to encode grammatical functions and relations.

Having said that, morphology is not entirely absent from DIP, and certain grammatical domains such as tense marking combine morphology and syntax in very unusual ways. This makes it difficult to reach a univocal classification of DIP in any morphological typology, but in fact it has already been recognised that categories such as 'isolating', 'agglutinative' or 'inflectional' (to use the terminology derived from Schleicher 1859) are to be seen as idealisations defining the extremes of continua. That is precisely the approach of Comrie (1989) as he proposes the indexes of synthesis and of fusion. Comrie's index of synthesis provides a tool to classify languages (and parts of grammatical systems) according to the number of morphemes per word, along a scale ranging from strictly 'isolating' (i.e., one morpheme per word) to highly 'polysynthetic' (i.e. many morphemes per word). The second of Comrie's indexes is that fusion, which measures the degree to which a single morpheme may express several distinct units of meaning. At the extremes of this continuum are highly 'fusional' (i.e., one morpheme tends to express more than one meaning) and highly 'agglutinative' (i.e., each unit of meaning tends to be expressed by a dissectable morpheme) languages. With regard to these criteria, then, DIP must be classified as a *highly isolating* and *mostly agglutinative* language. This classification respects the fact that most words in DIP are monomorphemic, while simultaneously revealing that morphology still plays a role in certain subdomains - most notably in the verbal domain (see section 6.2) - and also that a considerable amount of suppletion pervades the language's morphological paradigms - as for instance that of personal pronouns (see section 6.6.1).

6.1 Defining morphosyntactic categories

The question of which criteria to use in order to define PoS categories is far from resolved, and in fact there seems to be an element of arbitrariness to it that allows different scholars to pick and choose their own preferred methodologies.[1]

As one approaches the task, an important variable to take into account is whether to base one's PoS analysis on semantic or purely formal criteria, or otherwise. The traditional semantic approach which attempted to equate for example *nouns* with words referring to people, places or things (see Schachter 1985) has long come under attack - consider for instance the semantics of the English nouns *wrath* or *complementarity*. In tune with much of the modern PoS research, I adopt a rather morphosyntactic approach, considering syntactic distribution and the acceptability of certain morphological operations as valid classificatory variables. It would be rash, however, to completely disregard semantic classes and their interaction with PoS. Indeed, Wierzbicka (2000)) suggests that, given the heterogeneity of morphosyntactic structures worldwide (Haspelmath 2007b), semantics may be the sole factor on which to establish common ground for cross-linguistic comparisons. She therefore proposes that a highly restricted, seemingly universal set of basic semantic primitives (e.g. THINK, KNOW, WANT, FEEL, SEE and HEAR as representatives of 'Mental predicates') can be used for the delimitation of PoS systems and for their comparison across languages.

Another fundamental question is often whether to strive for a set of PoS categories which is exclusively language-specific or whether to bring in cross-linguistic information in an attempt at reflecting some degree of universality (e.g. cognitive universals or Universal Grammar). A balance needs to be struck between the two approaches. Cross-linguistic studies of PoS systems (which are at present not satisfactorily worked out) should in principle reveal patterns and prototypes capable of conducting the description of any language's PoS; but ultimately one must at least entertain the possibility that 'languages could differ from each other without limit and in unpredictable ways' (Joos 1957:96) and as a matter of course allow the language under analysis to construct its own PoS divisions and subdivisions. Some typologically-oriented theories have attempted to classify PoS according to their functional distribution, chief among which the typology of PoS systems often referred to as the Amsterdam Typology or the Amsterdam Model (see e.g. Hengeveld 1992b, Hengeveld et al. 2004), while Croft's universal-typological theory of PoS (see Croft 2000, 2003) proposes to map language-specific data onto a universal conceptual space. Both these theories are built on tried and tested concepts such as Predication, Reference and Modification and therefore provide solid ground to delimit the main PoS of a language.

The Amsterdam Model, the original purpose of which was to investigate the possibility of an implicational hierarchy shaping the PoS systems of the world's languages, considers four central functionally-defined PoS categories and investigates how different languages map onto it (i.e., it allows for different languages to not realise all the proposed distinctions

---

[1] According to some authors, the search for crosslinguistically comparable categories is even fundamentally flawed; Haspelmath (2007b) admits that there may be no way to attain such unversally applicable criteria at all, instead proposing that PoS categories are in the end strictly language-specific (see also Ansaldo et al. 2008).

or to do so in different ways). The four PoS are thus defined:

- Verbs: elements which, without further measures being taken[2], have a predicative use only;
- Nouns: elements which, without further measures being taken, can be used as heads of a referential phrase;
- Adjectives: elements which, without further measures being taken, can be employed as modifiers of a nominal head;
- Adverbs (restricted to traditional Manner Adverbs: elements which, without further measures being taken, can be employed as modifiers of a predicative head.

Given the degree of abstraction necessary for any model of PoS to be universally significant, both the Amsterdam Model and Croft's universal-typological theory of PoS are more useful to understand core categories - such as noun, verb or adjective - than relatively minor categories such as prepositions or interjections. It is to be expected that further research will extend the scope of cross-linguistically-valid models of PoS, but for the time being one needs alternative criteria to classify the lexemes not covered so far. The criteria that I base my PoS values on are spelt out at the opening of the relevant sections below.

Deciding on the nature of the criteria adhered to and the theoretical orientation does not exhaust the choices one has to make when describing a language's PoS. Morphological and distributional considerations do not always result in unequivocal category boundaries and automatic category-assignment; a particular lexeme may fulfill most of the requirements of a particular category but fail to acknowledge others, and in such cases it is a matter of choice whether to assign that element to this category or to create a separate (sub)category for similar lexemes. Faced with such dilemmas, scholars fall onto the categories of 'lumpers' (i.e., those who prefer macro-categories which do not require their members to behave entirely alike) or 'splitters' (i.e., those who would rather subdivide categories in pursuit of absolute homogeneity). One way to avoid splitting *ad aeternum* has been to apply the notion of 'prototype' to the definition of PoS. According to such approaches, PoS can be categorised around prototypes, and elements thus classified are allowed to depart to different degrees from it (Croft 2003). It is plain to see that such a concept tends particularly towards 'lumping' rather than 'splitting'. Several of the categories I propose are construed as prototypical, but I have felt the need to 'split' certain others (in particular those involving grammatical words) so as to gain in descriptive clarity.

6.2   Verbs

According to Hengeveld's (1992b) criterial test, verbs are elements which, without further measures being taken, are allowed only in predicative position. In DIP, a class of lexemes clearly satisfies this criterion: these elements can be clausal heads and determine the

---

[2] The phrase 'without further measures being taken' ensures that the proposed tests function only in the absence of any additional morphosyntactic means to indicate (non-prototypical) functions. For instance, although the English adjective *heavy* can be used predicatively as in the sentence *Lead is heavy*, the atypicality of this use is given away by the obligatoriness of an additional element, viz. the copula *is*.

argument structure of the entire clause. The simplest non-elyptical declarative clause in DIP consists minimally of an intransitive verbal head and one argument, so this type of construction is used in (19) as a test for the capacity of various forms to occupy a clause-heading position without further measures being taken (e.g. without the intervention of a copula). According to this criterion, *paro* 'stopped' emerges as a verbal form while *amig* 'friend' and *ispɛrt* 'smart' do not.

(19)   a.   *Fabian par-o.*
           F.   stop-PST
           'Fabian stopped.'

       b.   * *Fabian amig.*
           F.   friend

       c.   * *Fabian ispɛrt.*
           F.   clever

Another powerful indicator of verbness is inflectional morphology. In DIP, verbs are the only words that truly admit morphological inflection. DIP verbs are organised in inflectional classes - identified by different thematic vowels that attach to a verbal root -, which dictate the form of the inflectional suffixes to follow (see also Luís 2008). As an example, let us say that a regular verb of the -*a*- inflectional class such as *para* 'to stop' or *mara* 'to build' inflects according to the following paradigm:

a) Infinitive: root + theme + ø - e.g. *para, mara*
b) Non-Past: root + ø + ø - e.g. *par, mar*
c) Past: root + ø + o - e.g. *paro, maro*
d) Participle: root + theme + d - e.g. *parad, marad*

The inflectional paradigms for the remaining inflectional classes -*e*- and -*i*- are somewhat similar but with important differences. The paradigm for the -*e*- class, exemplified here with *sabe* 'to know' and *aprəse* 'to seem' is as follows:

a) Infinitive: root + theme + ø - e.g. *sabe, aprəse*
b) Non-Past: root + ø + ø - e.g. *sab, aprɛs*[3]
c) Past: root + theme + w - e.g. *sabew, aprəsew*
d) Participle: root + i + d - e.g. *sabid, aprəsid*[4]

The final regular inflectional class is that in -*i*-, exemplified with the verbs *pidi* 'to demand' and *uvi* 'to hear'. The following general rules apply:

---

[3]Given their syllabic structure, stress falls on the last syllable of the root in non-Past verbal forms. As expected, the phonological rules governing schwa-deletion (see section 5.4) and the dispreference of schwa in stressed position alters the final root vowel, resulting in -ɛ- in the case of *aprɛs* 'seem(s)'.

[4]The form *aprəsid* does not actually occur in the corpus, which is understandable given how marginal the occurrence of both participle forms and the verb *aprəse* is. The non-Past form was reconstructed according to rather regular morphological rules so as to complete the paradigm.

a) Infinitive: root + theme + ø - e.g. *pidi*, *uvi*
b) Non-Past: root + alternating vowel + ø - e.g. *pɛd*, *ov*[5]
c) Past: root + theme + w - e.g. *pidiw*, *uviw*
d) Participle: root + i + d - e.g. *pidid*, *uvid*

Not all verbs fall onto these predictable inflectional paradigms. A number of high-frequency verbs such as *vay/ir* 'to go', *vi* 'to come' or *da* 'to give' inflect irregularly. In the case of *vay/ir*, the paradigm is entirely suppletive, as shown below:

a) Infinitive: *vay*[6]/*ir*
b) Non-Past: *vay*
c) Past: *foy*
d) Participle: *id*

A conspicuous Past suffix which is unrepresented in the examples above is -*iŋ*, occurring amply in the form *tiŋ* 'had' but also in *kiriŋ/keriŋ* 'wanted', *pudiŋ* 'could' and *sabiŋ* 'knew'.[7]

Most DIP lexemes resist any kind of morphological segmentation, let alone inflection; those fall outside the verb category. DIP distinguishes two copulas, both of which inflect to show tense and aspectual information: the stage-level copula *te* (20) and the individual-level copula *ɛ* (21):

(20)   a.   *yo tə   duẽt.*
               1s COP$_s$.NPST ill
               'I am ill.'

           b.   *yo t-iŋ   duẽt.*
               1s COP$_s$-PST ill
               'I was ill.'

---

[5] As these examples clarify, the type of thematic vowel alternation in class -*i*- verbs is not entirely dependent on phonological operations; there is no phonological constraint on the sequences */pid/ and */uv/. In fact, these formal alternations in the DIP present forms reflect very closely the Standard Portuguese forms they stem from: in this case, SP *pede* 'asks for' and *ouve* 'hears' - notice that the etymological reference for DIP present forms usually corresponds to the present third person singular form in Standard Portuguese. From the perspective of a DIP speaker, though, these vowel alternations are probably best interpreted as suppletive.

[6] *Vay* is exceptional among other DIP verbs in that its infinitival form is not derived from a Ptg. infinitive but rather a finite form (Ptg. *vai* 'go.PRES.IND.3s'), although a less frequent form *ir* is also attested. The same holds true for cognate verbs in other Portuguese-lexified Creoles, both in Africa and in Asia, which prompted Clements (2000) to admit some degree of carry-over between the Portuguese Pidgins of West Africa and their Asian counterparts.

[7] There is reason to expect this morpheme to have extended from *tiŋ* towards the other verbs, as this is the only form for which the Portuguese cognate contains a comparable suffix: Pt. *tinha* 'have.PST.IMP.1/3s'. By falling outside the regular inflectional paradigms of Portuguese and including the phoneme /i/ in the 1/3s form of the Past imperfective, the translation equivalent of 'to want', 'can' and 'to know' must have been particularly receptive to the application of this Past morpheme.

(21) a. *es* *mĩ* *kubɛrt* *ɛ.*
DEM$_p$ 1s.POSS blanket COP$_i$.NPST
'This is my blanket.'

b. *es* *mĩ* *kubɛrt* *ɛ-r.*
DEM$_p$ 1s.POSS blanket COP$_i$-PST
'This was my blanket.'

For a full account of non-verbal predication , see 7.1.3. Likewise, there is a grammatical verb *tə/tiŋ* which functions as an auxiliary (22), calling onto itself much of the tense-aspect-mood information necessary for the predication.

(22) a. *yo tə* *brĩk-a* *kaz* *ko* *saykəl.*
1s IPFV.NPST play-INF house INS bicycle
'I am playing with my bicycle at home.'

b. *yo t-iŋ* *brĩka* *kaz* *ko* *saykəl.*
1s IPFV-PST play.INF house INS bicycle
'I was playing with my bicycle at home.'

Some Tense/Aspect/Mood (henceforth TAM) markers, such as for example the auxiliary use of the *kere* 'must', are also verbal in nature (see 7.1.2); others, including the irrealis marker *a(d)* fail to be classified as verbs given that they do not inflect.

Reduplicated verbs do not occur very frequently, but when they do one can abstract intensified semantics, as in (23):

(23) *si yo a* *bebe∼bebe* *tə* *durmi.*
if 1s IRR.NPST drink∼drink.INF IPFV.NPST sleep.INF
'If I were to drink a lot, I would [always] fall asleep.'

On a final note, it is interesting to notice that one of the traditional defining characteristics of the class of verbs, viz. the fact that they are an *open class*, is not entirely straightforward in the case in DIP. On the one hand, the language does have a (socially-motivated) predisposition to borrow verbs from Ptg., but on the other hand there is no morphological means of deriving denominal or deadjectival verbs. In addition, I have not observed any strategy for the morphological incorporation of Indic verbs of the type described for Korlai Indo-Portuguese - which amounts to inserting Indic verbal morphemes into a separate conjugational class (Clements 1996, Luís 2008). Instead, DIP employs an analytical strategy known in the literature on Indo-Aryan languages as 'conjunct verbs' (e.g. Masica 1991:368*ff*), which consists of creating a complex predicate through the combination of a nominal or adjectival preverb and a light verb, in this case the verb *faze* 'to make' - see section 7.1.4 for a full description of the construction.

6.3 Adverbs

Adverbs are usually understood to be optional modifiers of non-nominal arguments (see Van der Auwera 1998a). As a result of such a broad definition, the notion of 'adverb' usually encompasses elements with strikingly different semantics and can only be defined by looking at their distribution within the clause. Yet, adverbial functions can be taken on by several constructs, such as prepositional phrases, adjunct clauses or noun phrases, and as such the task at hand is to ascertain whether any lexemes can be used in such functions 'without further measures being taken' (Hengeveld 1992b).

6.3.1 Manner Adverbs

Out of the traditional semantic subdivisions of adverbs, Hengeveld (1992b) as well as Sasse (1993:quoted in Wierzbicka 2000) and Ramat and Ricca (1994) take Manner Adverbs to be central in defining whether or not a separate class of adverbs can be posited for a particular language. In modern DIP, there is no productive derivational strategy to construct Manner Adverbs, but there is a small subset of manner-defining elements which are only allowed as modifiers of the predication. Those include *asĩ* 'thus', *bẽy*[8] 'well', *(də)fors* 'fast', *dret* 'right, correctly', *jũt* 'together, along', *lɔg* 'immediately, early', *mal* 'badly', *suziŋ* 'alone', and *vagar* 'slowly, low [volume]'. An example with *bẽy* is given in (24):

(24)   *kɔm   bẽy i   vẽy.*
        eat.NPST well and come.NPST
        'Eat well and (then) come.'

These lexemes are only allowed as modifiers of the predicate and not as noun modifiers, which establishes a solid divide between adverbs and adjectives. It is, nonetheless, possible to detect some similarities between adverbs and adjectives in DIP, chief among which the fact that both can be modified by intensifiers. In addition, there are scattered examples of what might be called zero-derivation, in which a prototypically adjectival element modifies a verbal predicate. One case in point is that given in (25), where the adjective *bunit* 'beautiful' functions as a manner adverb equivalent to *bẽy*:

(25)   *dɔn   Angelina kuziŋ   bunit   nə?*
        madam A.   cook.NPST beautiful REQ
        'Madam Angelina cooks well [lit. beautiful], doesn't she?'

Section 7.3 describes the paralell structures of Adverb Phrases (AdvP) and Adjective Phrases (AdjP), drawing attention to their similarities and differences.

---

[8]As a degree indicator, *bẽy* also modifies adjectives (see section 7.3).

6.3.2 Spatial Adverbs

Another important set of adverbs indicate location and direction, and in order to capture both readings they will be termed Spatial Adverbs. These include *(a)bayx* 'below', *aki* 'here', *ali* 'there', *dẽt* 'inside', *fɔr* 'outside', *frẽt* 'ahead', *imɔr* 'away', *la* 'there', *lõj* 'far', *pɛrt* 'near', *(i)sim* 'above' and *(a)tras* 'behind'. As demonstrated in (26), DIP makes no formal distinction between allative and stative predication, with Spatial Adverbs usually indicating direction when accompanied by a motion verb such as *vay/ir* 'to go' and location in any other case.

(26) a. *yo foy   frẽt.*
           1s go.PST ahead
           'I went ahead/to the front.'

     b. *yo sẽt-o   frẽt.*
           1s sit-PST ahead
           'I sat at the front.'

6.3.3 Temporal Adverbs

Temporal Adverbs locate the clause in time and, like Spatial Adverbs, are largely deictic. The DIP forms are *agɔ(r)* 'now', *amiŋã* 'tomorrow' , *ãt* 'before', *dəpəy* 'afterwards', *oj* 'today', *õt* 'yesterday', *pimer* 'earlier', *(sə)sed*[9] 'early' and *tard* 'late, afternoon'. The words *maŋã* 'in the morning' and *noyt* 'at night', despite doubling as nouns, often occur adverbially without any markers of non-prototypicality, but often they are preceded by the preposition *də* (cf. 6.7).

6.3.4 Phasal Adverbs

Phasal Adverbs indicate the actuality and/or (dis)continuity of a particular state. According to Van der Auwera (1998b), such adverbials are used to express inchoative (e.g. *already*), continuative (e.g. *still*), discontinuative (e.g. *no longer*) and continuative negative (e.g. *not yet*) semantics. DIP has dedicated adverbs for the inchoative and continuative only, *jə*[10] and *ain* respectively. These are exemplified in (27):

---

[9]Some adult informants single out the expression *maŋã səsed* 'early in the morning' as one of the characteristics of DIP as spoken by the generation of their parents. This brings to mind the possessive constructions of other IP varieties (such as e.g Korlai IP or Cannanore IP) which follow the structure Possessor-*Gen*-Possessee: *maŋã su sed = morning GEN early(time). If this interpretation is correct, one can entertain two diverse diachronic explanations: either the crystallised expression reflects a past stage of DIP, or it may have been borrowed in its entirety from other varieties of IP.

[10]In addition, *jə* also marks an action as occurring immediately upon a given event. In this sense, it is very close to the adverb *lɔg*, see 6.3.1.

(27)   a.   *exam jə*   *kab-o.*
              exam already finish-PST
              'The exam is already over.'

      b.   *el ain te*   *atras.*
              3s still EXS.NPST back
              'He's still back there.'

The corpus does contain a continuative negative construction, consisting of the negated continuative adverb (28a), and a discontinuative construction, which employs a negated inchoative adverb (28b):

(28)   a.   *ɔr*   *nə*   *fik-o*   *ain.*
              hour NEG$_{cl}$ become-PST still
              'It's not time yet.'

      b.   *yo jə*   *nə*   *te*   *duẽt.*
              1s already NEG$_{cl}$ COP$_s$.NPST ill
              'I'm no longer ill.'

6.3.5   Frequency Adverbs

Similarly to Phasal Adverbs, Frequency Adverbs are often instrumental in expressing aspectual information. In DIP, the relevant distinction is between a negative frequency adverb *nuk*[11] 'never' and the positive frequency adverb *sẽp*, which conflates the meanings of 'always' and 'often' or 'a lot'. More detailed frequency information is provided by adjuncts such as *bastãt vez* 'many occasions'. The case of the adverb *utrez* 'again' may raise questions as to its internal structure, given that its provenance from the compound *otr vez* 'another occasion' is immediately clear. It is the frequent phonological erosion of the erstwhile compound that suggests a process of grammaticalisation has given rise to a monomorphemic adverb. The same reasoning applies to the form *tudi* 'everyday', from the expression *tud di* 'every day'.

6.3.6   Sentence Adverbs

Sentence adverbs are defined as adverbs with scope over the entire proposition instead of just the predice (such as for instance *apparently* in *She is too tired, apparently*). These adverbs typically convey modal information of various kinds.

---

[11] Like the universal negative quantifiers *nad* 'nothing' and *nĩge* 'nobody', *nuk* triggers negative concord on a syntactic level; see section 7.7.

The corpus contains a Sentence Adverb denoting weak probability, viz. *pədser* 'perhaps'. The issue might arise whether *pədser* is a syntactic construct rather than a monomorphemic adverb, spurred by the fact that the auxiliary verb *pɔd* 'can' often expresses possibility (as well as ability) in constructions in which it takes a non-finite verbal complement. Recall, for the discussion, that Ptg. *poder* 'can' behaves similarly, and in fact *pədser* is clearly derived from the Ptg. expression *pode ser* 'can be'. As far as DIP *pədser* is concerned, however, I would argue that we find hints of lexicalisation in that a) *ser* is not a productive verbal form in DIP; b) while expressions with *pɔd* + v.INF maintain stress on both elements, the first syllable in *pədser* is predictably unstressed, and the vowel reduced to schwa; c) the occurrence of this element within clauses with a distinct verbal head and no indication of either coordination or subordination (see 29a); and d) there is no record of the putative *pɔd* in *pədser* being inflected.

The word *parɛs* 'apparently' also belongs to the category of Sentence Adverbs, denoting weak indirect knowledge of a particular proposition. Its use is exemplified in (29b).

(29) a. *pədser yo vay   Una amiŋã.*
           perhaps 1s go.NPST Una tomorrow
           'Maybe I will go to Una tomorrow.'

      b. *parɛs   use t-iŋ   nə Una õt.*
           apparently 2   EXS-PST LOC U. yesterday
           'It seems you were in Una yesterday.'

6.3.7   Adverb reduplication

Adverb reduplication[12] is not uncommon in DIP. In such cases, intensification is the only reading available, as exemplified in (30):

(30) a. *ãta nɔs fal   pok mays miyɔr asĩ, ẽ kaz   nɔs fal*
           then 1p speak.NPST little more better thus in house 1p speak.NPST
           rough~rough.
           rough~rough
           'Then we speak a little bit better, at home we speak very rough.'

---

[12] Adverbial phrases, rather than monomorphemic adverbs, are similarly reduplicated, as shown in the following example:

(1) *yo tə   ka-i   aŋũ vez   aŋũ vez.*
     1s IPFV.NPST fall.INF sometimes sometimes
     'I fall down sometimes, sometimes.'

b. *nɔs fal*   *lɔg~lɔg,*   *lɔg~lɔg*
1p   speak.NPST immediately~immediately immediately~immediately
purtəgez   *a sĩ.*
portuguese thus
'We speak Portuguese very quickly, very quickly like this.'

This fact establishes an important connection with adjectives, the reduplication of which also carries intensifying semantics - see section 6.5.1 and also 7.3 for further information and examples.

6.4   Nouns

The category of nouns, readily observable in several European languages because of certain morphological idiosyncrasies, does not always translate easily onto the PoS systems of other languages, in particular in highly isolating languages. The most problematic boundary is often that between nouns and adjectives. As expected, the division between DIP nouns and adjectives cannot primarily be founded on morphological considerations, as the elements of both these categories are formally highly stable and invariant. Apart from the obvious but elusive semantic criteria that are normally adduced to justify the noun/adjective divide, it is their distinct distributional characteristics which allow us to tease them apart. In DIP, not all content words are allowed, without any further measures being taken, as heads of referential phrases which take modifiers, so we can take this as an initial criterion to delimit our concept of the prototypical noun. In the following examples, the lexeme *bud* 'ghost' is the head of the noun phrase, functioning as a subject-like argument to the intransitive verb *sai* 'come out'; notice how the sentence in (31a) can be reduced to (31b) but not (31c):

(31)   a.   *ikəl   doy gran bud   a   sa-i.*
          DEM_d two big   ghost IRR.NPST leave-INF
          'The two big ghosts will come out.'

      b.   *ikəl   doy bud   a   sa-i.*
          DEM_d two ghost IRR.NPST leave-INF
          'The two ghosts will come out.'

      c. * *ikəl   doy gran a   sa-i.*
          DEM_d two big   IRR.NPST leave-INF

Adjectives, on the other hand, will be defined below as lexical (or content) words which modify referential heads. Still, I do not wish to enforce too strict a distinction between these two categories, preferring to treat them as prototypes at the extremes of a scale of referentiality which allows some mobility. One of the reasons for this is that prototypical nouns can on occasion modify other nouns, as in (32):[13]

---

[13]The issue of whether this is an instance of an adjectivised noun or rather appositional modification is unclear, and possibly theory-dependent.

(32) *ali*   *nɔs vi-w*   Blackmail *sinem.*
      there 1p   see-PST Blackmail film
      'There, we saw the film *Blackmail*.'

Bearing in mind the gradience of the noun-adjective cline, we can now proceed with a definition of the prototypical noun. Following Hengeveld's (1992b) proposal, I take a word's capacity to head a referential phrase - without resort to any morphosyntactic indicators of non-prototypical function - to be a sufficient test for noun categorisation. Adding any further distributional criteria would result in unnecessary categorical 'splitting'; different nouns may behave differently with regard to modification, for instance, but in most cases these differences boil down to semantic considerations. It is therefore interesting to explore to what extent one can recognise subgroupings within the category of nouns in DIP, and how semantics constrains the distributional behaviour of the different nouns.

In DIP, there is no formal distinction between what is usually termed proper names and other nouns, with the exception of those structural traits expected from a maximally unambiguous and definite referent: a certain difficulty in accepting modification, in particular by relative clauses and quantifiers.

Similar semantic considerations can be adduced to understand the marginal distributional differences between count nouns and mass nouns. Count nouns such as *chumas* 'pillow' admit numerals more readily than mass nouns such as *arey* 'sand', but this is where the formal differences end; unlike in many other languages, no split exists between quantifiers which modify mass nouns and those reserved for count nouns, so that for instance *bastãt* 'much/many' or *aŋũ* 'some' can be used with both in exactly the same way. Compare the constructions in (33), where they modify a count noun, with those in (34), in which the head of the NP is a mass noun[14]:

(33)   a.   *sey chumas*
            six pillow
            'six pillows'

       b.   *bastãt / aŋũ   chumas*
            much   some pillow
            'many/some pillows'

(34)   a.   ? *sey arey*
              six sand
            'six sands'

---

[14]This fact is not entirely surprising given that Ptg., the main lexifier of DIP, also uses the cognate quantifiers with both count and mass nouns. The difference is that in Ptg. the quantifiers inflect for number (and gender) in agreement with the head of the NP to the effect that they always appear in their singular form with mass nouns and the plural form with count nouns: e.g. *alguma-s almofadas* 'some pillows' vs. *alguma-Ø areia* 'some sand'.

118

b. *bastãt / aŋ̃ũ arey*
   much   some sand
   'much/some sand'

The fact that DIP does not distinguish mass vs. count noun quantification is an interesting typological trait. Rijkhoff (2000, 2002, 2008) provides a crosslinguistic typology of nouns according to their referential characteristics, based on the features SHAPE (i.e., whether or not the referent is mentally construed as having set boundaries and therefore being countable) and HOMOGENEITY (i.e. whether the basic referent is construed as agglomerative or singulative). To follow his terminology, DIP nouns fall squarely into the category of *set nouns*, meaning that the referents are construed as having set boundaries (+ SHAPE) but left unspecified with regard to HOMOGENEITY. In practical terms, this means both that the shape of DIP bare nouns indicates nothing in the way of number (35a), and that DIP nouns can enter a direct relationship with (numeral) quantifiers without the intervention of classifiers (35b). The number ambiguity is dealt with by means of a collectivising marker which does not co-occur with a numeral. The collectiviser may be inflexional in some languages with set nouns, but in the case of DIP it is taken over by the word *tud*, which is also a quantifier meaning 'all' (35c):

(35)   a.   *adiw*
           fox
           'fox / foxes'

       b.   *sey /   bastãt adiw*
           six many fox
           'six / many foxes'

       c.   *ikəl   tud   Ø / * sey adiw*
           DEM$_d$ COLL Ø   six fox
           'the foxes'

For a more concrete account of the role and distribution of the collectiviser, see 7.2.7.

The previous examples demonstrate that no productive nominal inflection exists in DIP to express those grammatical functions that other languages code morphologically (e.g. number, class[15], definiteness), with the exception of the rather marginal use of nominal reduplication as a pluralising strategy (for which, see section 7.2.7.1). There are also remnants of an etymological diminutive suffix *-iŋ* in words such as *ratiŋ* and *avziŋ* 'grandmother', but this is no longer productive (see 8.2.1). For instances of derivational morphology in DIP, refer to the sections dealing with the language's lexicon (chapter 8).

---

[15]Gender is not morphologically marked, and it is not a relevant category for most nouns. However, there is a subset of (typically human) nouns which establish gender oppositions on a lexical level, corresponding to natural gender - see 8.6.1.

Nouns can be reduplicated in DIP, although this occurs very rarely. The most straightforward semantic value of nominal redulication is distributive. Consider the sentence in (36):

(36) *es*   *tud rəkri*   *tud ε*   *də moyr∼moyr.*
      DEM$_p$ all food.stall all COP$_i$ of muslim∼muslim
      'All these food stalls belong to (various) Muslims.'

6.5   Noun modifiers

Within the NP, the head noun can be modified by elements of several orders, and setting categorial boundaries between them is often challenging. Distributive constraints are good initial criteria, and therefore I must begin by succinctly describing the structure of the language's NP. The following figure is a simplified schematic representation of the NP with the various positions of single-word noun modifiers identified by numbers, for the time being. Every modifier position is only optionally filled, and the actual realisation of such long strings of modifiers is extremely rare.

(37) $3 + 2_a/2_b + [\text{ X} + 1] + \text{N} + \text{Y}$

In (37), the elements represented by a letter are not relevant for our present purposes: *N* stands for the head noun, *X* is a modifier of *1* rather than *N*, and position *Y* can only be occupied by phrasal modifiers. This section is therefore concerned with the elements able to occupy positions *1* through *3*, which will be discussed one by one. For the moment, I will simply flesh out the scheme with the modifier categories I will subsequently propose: adjectives (optionally modified by an intensifier, *X* in this scheme) filling position *1*, quantifiers or ordinals in position *2* (see 6.5.3 for a justification of the dichotomy), and deictic modifiers (both possessives and demonstratives) in position *3*. Position *Y* is reserved for prepositional phrases and relative clauses, which often take over the functions of some of the single-word modifiers. The scheme can now be filled in as in (38).

(38)   Dct + Qtf/Ord + [Int + Adj] + N + PP/Rel

The various categories proposed here are constructed so as to reflect these distributional differences, although there is an element of arbitrariness in the definition of PoS systems. That is particularly evident if the elements under analysis present strong functional similarities - as in the case of the various noun modifiers - and when morphology offers little in the way of distinguishing the various elements - which is the case in highly isolating languages such as DIP. As a result, categorial boundaries must therefore be allowed a certain dose of fuzziness.

Some interrogative words typically occur in modifier position. In view of that, I will add a section dealing with what, reflecting both their syntactic position and their function,

will be termed *adnominal interrogatives*.

6.5.1 Adjectives

The class of adjectives is hardly ever a completely straightforward category in the world's languages. Much of the theoretical debate actually concerns their formal similarities with either verbs or nouns given that, in various languages, the translation equivalents of English 'yellow' and 'weak', for instance, behave formally more like verbs or like nouns. Conveniently, Hengeveld's typology provides useful formal criteria to distinguish these three word classes; in this framework, adjectives are defined as lexical elements allowed as modifiers of referential heads, without any further measures being taken. In order to set the boundaries between adjectives and the remaining noun modifiers, we will need to appeal not only to distribution and co-occurrence but also to the lexical/grammatical word dichotomy, in which adjectives are seen to represent the former.

The criteria so far adduced refer to the attributive use of adjectives, but these elements can also be used predicatively, and in DIP they do so more often than not. In this position, certain adjectives may select both copulas, whereas others show clear preference for the individual-level copula. The crucial difference here is that of perceived time-stability, with the individual-level copula ε introducing more or less permanent states while the stage-level copula *te* is reserved for more temporary or recent conditions (see section 7.1.3). Let us illustrate this by referring to the varying behaviour of two adjectives, *bõ* 'good' and *nov*[16] 'new, young' when combined with the stage-level copula. It is clear that, from a semantic point of view, *nov* is more time-stable than *bõ* in that quality is more variable than age; in this sense, it is understandable that *bõ* is more acceptable in combination with copula *te* than *nov*. For comparative reasons, I include an example of a noun (*pulis* 'policeman') in predicative function, as one would expect a noun typically to denote a rather time-stable referent. The first sentence in (39) is the standard way of indicating that someone is in good health or faring well generally, while the second might be interpreted as an indication that F. is looking younger than usual. The final sentence, with *pulis*, cannot be an instance of a predicative nominal - with nouns, the verb *te* has a possessive reading only (see section 7.1.3.4).

(39) a. *Fabian te   bõ.*
          F.   COP$_s$.NPST good
          'Fabian is fine.'

      b. ? *Fabian te   nov.*
           F.   COP$_s$.NPST young
          'Fabian is/looks young.'

---

[16] These particular adjectives were selected in accordance with Dixon's (1982) suggestion that, cross-linguistically, value and age are among the semantic types most likely to be expressed by an adjective; Wierzbicka (2000) partly echoes this intuition when she proposes GOOD to be one of the semantic primes associated with the word class of adjectives.

c. * *Fabian te*   *pulis.*
     F. COP_s.NPST policeman

Another conspicuous characteristic of prototypical adjectives is their gradability[17], particularly in predicative position. This is demonstrated for *vɛy* in (40), in which the adjective is modified by an intensifier *bẽy* 'very' and *pok* 'little':

(40)   *bẽy* / *pok*   *vɛy* *ɛ.*
       very   little old COP_i.NPST
       '(He) is very/a bit old.'

This observation indicates that an adjective can head an AdjP in which it is modified by an intensifier (as described in 7.3). However, gradability, as a logical concept, is still gradient. Notice, in (41), the modification of the term *ĩdu*, a term which is quite prototypically noun-like:

(41)   *el* *ɛ*   *ĩdu,*   uncle. uncle, *el* *ɛ*   *bẽy* *ĩdu.*
       3s COP_i.NPST hindu uncle   uncle   3s COP_i.NPST very hindu
       'He's Hindu, uncle. Uncle, he's very Hindu.'

DIP makes regular use of reduplication, the most pervasive type of which involves adjectives.

(42)   a.   *ikəl*   *fin~fin*   *pɛd.*
           DEM_d small~small stone
           'That very small stone.'

       b.   *pok~pok*   *sɔ*   *sab.*
           little~little only know.NPST
           '(I) only know a little bit.'

As shown in (42), adjective reduplication has exclusively an intensifying function. DIP adjectives do not inflect in any way other than total reduplication.

6.5.2   Quantifiers

The category of quantifiers includes both cardinal numerals and quantifiers proper (i.e., non-numeral quantifiers), given that in DIP there is little distributional distinction between the two within the NP (see section 6.5.3 to understand why ordinal numerals are left out

---

[17]See Bhat et al. (2000) for a crosslinguistic account of the phenomenon.

from this count). They are mutually exclusive within the NP, except if the quantifier is an intensifier embedded in the AdjP (see 7.3). The prototypical function of quantifiers is that of noun modification, as shown in (43) for *tud* 'all' and *doz* 'twelve'. Consider the use of both the quantifier *tud* 'all' and the numeral *doz* 'twelve' in the following example:

(43) a. *nə   mūbəy   tud jardĩ   ɛ   gran.*
          LOC Mumbai all garden COP$_i$.NPST big
          'In Mumbai, all gardens are big.'

      b. *Lisa te   doz   gat dẽt   d-ɛl kaz.*
          L.   have.NPST twelve cat inside of-3f house
          'Lisa has twelve cats at home.'

For a more complete description of quantifiers as nominal modifiers, see section 7.2.2. In addition, widespread ellipsis dictates that quantifiers can also occur in seemingly pronominal contexts, in particular whenever the referent is known or assumed. Consider the use of both the non-numeral quantifier *tud* 'all' and the numeral *ũ* 'one' in the following example:

(44) a. *tud te   ku saykəl i   ũ   te   ku TV   sɔ.*
          all IPFV.NPST INS bicycle and one IPFV.NPST INS television only
          'Everyone played with their bicycles and only one with the TV.'

      b. *pok sɔ   sab.*
          little only know.NPST
          '(He) only knows a little.'

Pok 'little', used in (44b), is particularly common on its own, and it often occurs in a (quasi-)nominal construction with the indefinite marker *ũ*. It appears that this element is not just a quantifier but also a lexical noun referring to 'a portion'. Consider for instance the following example, in which *pok* is modified by the NP-negator *niŋũ* 'no [X]':

(45) *nã   apaŋ tẽp, niŋũ   pok.*
      NEG$_{cl}$ find   time NEG$_{ct}$ little
      '(I) do not find the time, not even a little.'

Reduplicated quantifiers have a distributive function, as exemplified in (46):

(46) *tud   di   vay   da   trĩ~trĩt   rupi.*
      every day go.NPST give.INF thirty~thirty rupee
      'Every day (I) give them thirty rupees each.'

123

In this respect, quantifiers differ substantially from adjectives given that, as explained in 6.5.1, adjectival reduplication has intensifying rather than distributive semantics.

A complete description of cardinal numerals in DIP is given in 8.4.1. Non-numeral quantifiers show certain patterns of multicategoriality which set them apart from numerals. They are common agents of noun modification, and therefore must be discussed along with numerals, but many also transcend that function to occur as verb modifiers (adverbs) and operators within both the AdjP and the AdvP (intensifiers). These various functions are exemplified in the following examples. In (47), *muyt* is in turn a quantifier ((47a); notice also the presence of the comparative quantifier *men* 'less'), an intensifier modifying an adjective (47b), and an adverb (47c):

(47)   a.   *ali   muyt baruy, aki   men baruy.*
           there much noise   here less noise
           '[There is] much noise there, [there is] less noise here.'

       b.   *Leslie ε   muyt ispεrt.*
           L.   COP$_i$.NPST very smart
           'Leslie is very smart.'

       c.   *yo gost-o   muyt.*
           1s like-PST much
           'I liked [it] very much.'

Non-numeral quantifiers which also occur as intensifiers[18] include the mid-scalar quantifiers[19] *muyt* 'very/many/much', *bastãt* 'very/many/much' and *pok/pokxiŋ* '(a) little', as well as the comparative quantifiers *tãt* 'so (much/many)', *may(s)* 'more' and *men* 'less'.

In contrast, the existential quantifier *aŋũ* 'some', the NP negator *niŋũ* 'no [X]' (see 6.6.5 for the role of these quantifiers in constructing indefinite reference), as well as the universal quantifiers *ĩter* 'whole' and *tud* 'all' cannot operate within AdjP and AdvP. In particular syntactic contexts, *tud* also functions as a collectiviser (see 6.4 and 7.2.7).

6.5.3   Ordinals

Ordinals are formally linked to numerals, and they are mutually exclusive in the corpus. However, their contributions to the NP are starkly different. The function of ordinals, unlike that of cardinals, is not one of quantification but rather that of localising the referent

---

[18]No separate category of intensifiers is proposed for DIP precisely because none of the elements that take up this function does so exclusively. In addition to the quantifiers described here, the manner adverb *bɛ̃y* 'very' can also modify an adjective or another adverb. The syntax of both the AdjP and AdvP is described in detail in 7.3.

[19]I use the term *mid-scalar quantifier* as defined in Haspelmath (1997), in which it refers to quantifiers not in the extremes of a logical scale of quantification connecting the universal quantifier (e.g. Eng. *all*) to the universal negative quantifier (e.g. Eng. *none*).

within a sequence. One might wish to classify them alongside adjectives, but they would turn out to be rather atypical in that they cannot head an AdjP. For all these reasons, I have opted for classifying ordinals separately.

For the constituency of this category, described in detail in terms of their frequency as well as their morphological characteristics, see section 8.4.2.

6.5.4   Deictic modifiers

Deictic modifiers are those which prototypically ground the reference in reality at the moment of the utterance. In DIP these include both possessives and demonstratives, two semantically distinguished types of modifiers which, however, tend not to co-occur within the NP. Apart from their indexing potential, deictic modifiers may take up other functions, such as anaphoric reference, which operates within discourse alone. This is particularly evident in the case of demonstratives, as explained below.

One of the characteristics of deictics is that, apart from their adnominal function, they can also easily constitute a referential phrase of their own. The discussion must therefore be linked with that of personal pronouns (section 6.6.1) and demonstrative pronouns (section 6.6.2). For a syntactic account of deictic modifiers, see also section 7.2.3.

Possession may be indicated by a prepositional phrase embedded in the NP (see 7.2.3) or by adnominal possessives, non-lexical operators whose referent is retrieved either deictically from the speech-act situation or anaphorically from the discourse context. The paradigm of these forms is very close to that of personal pronouns, so much so that it is possible to argue that in most cases we are dealing here with specific genitive-marked personal pronouns. This paradigm contains monomorphemic forms only for 1s (equivalent to the oblique form of the 1s personal pronoun) and 1p (equivalent to the personal pronoun); the remaining forms consist of a contracted form of the preposition *də* 'of' plus the corresponding personal pronoun (cf. 6.6.1). The internal morphological structure of these forms is open to debate, especially given the fact that the preposition *də* is known to appear in its contracted form preceding other elements such as demonstratives or adverbs (see 6.7 and 7.4.3). The complete paradigm is given in Table 6.1 below:

	Singular	Plural
1st person	mi	nɔs
2nd person	duse	dusez
3rd person	del[m.], dɛl[f.]	de(l)z

Table 6.1: DIP possessive paradigm

DIP demonstratives reflect a two-tiered deictic space: proximal (*es* 'this', glossed 'DEM$_p$') and distal (*ikəl* 'that', glossed 'DEM$_d$'). These are widely used as deictic locators, as exemplified in (48):

125

(48) *aki aki, nə es igrej.*
      here here LOC DEM$_p$ church
      'Here here, in this church.'

Adnominal demonstratives are also instrumental in maintaining textual cohesion, as they recuperate previous topics through anaphoric reference. Consider the following example:

(49) *yo larg-o kurəsãw i vey. dəpəy* crocodile *foy ali kaz də*
      1s drop-PST heart   and come.PST then crocodile go.PST there house of
      irmã, foy ver pu ikəl kurəsãw.
      sister go.PST see.INF DAT DEM$_d$ heart
      'I dropped my heart and came here. Then the crocodile went to the sister's house, he went to check on the heart.'

The function of demonstratives therefore transcends deixis to mark definite reference (for which, see section 7.2.6) and also to aid discourse. For the pronominal use of demonstratives, refer to section 6.6.2.

Ot(r) 'other' is an absentive demonstrative in the sense that it signals the disjunction between a given referent and what can be apprehended from the deictic space (50a) and/or discourse context (50b):

(50)   a. *mĩ irmãw te nə ot∼ot iskɔl.*
            1s.POSS brother EXS.NPST LOC other∼other school
            'My brother is at a different school.'

        b. *də Go nã mãd-o purtəgal, ot lad mem dix-o.*
            from Goa NEG$_{cl}$ send-PST Portugal other place EMPH leave-PST
            '[They] didn't send [it] from Goa to Portugal, [they] left it some other place.'

Deictics *es*, *ikəl* and *ot(r)* have very peculiar semantics when applied in temporal phrases, for which see section 8.5.1.

6.5.5 Adnominal interrogatives

The question words *kwõt* 'how much/many', *kwɔl* 'what, which' and occasionally also *ki* 'what' differ from other interrogative words in that they occur within the NP, preceding the head noun whenever it is overtly expressed, and they are not pro-forms in any way. The following sentence exemplifies the use of the adnominal interrogative *kwõt*:

(51)   *kwõt*   *ɔr*   *jə*   *fik-oʔ*
       how.much hour already become-PST
       'What time is it?' [lit. 'How many hours']

No other NP modifiers co-occur with adnominal interrogatives. For a full account of the issue, see section 7.2.4.

6.6   Pro-forms

Pro-forms are defined in this study as elements which can occur as substitutes of entire clausal constituents such as NPs and adverbials, phrasal elements such as quantifiers, or as indexes of speech-act participants. Pro-forms are usually subcategorised according to the function of the element they replace or, to be more precise, the word-class of the element capable of fulfilling that function. It is therefore customary to speak of pronouns to refer to elements which stand for a noun (i.e, an NP; a clause argument), of pro-adverbs when the element being replaced is an adverbial, and pro-determiners as words that stand for or elicit a determiner. Despite these significant differences, it makes sense to discuss pro-forms together because elements from across these categories often share substantial functional and distributional characteristics - interrogative pro-forms and indefinite pro-forms are clear illustrations of this point. I will begin by discussing the subclasses which are entirely populated by pronouns (personal pronouns and demonstrative pronouns) before moving on to those that cut across the spectrum of pro-forms.

6.6.1   Personal pronouns

The general definition of a pro-form given above, viz. an element which substitutes for something else, needs to be re-worked when dealing with pronouns. It has already been pointed out (e.g. Evans 2000, Bhat 2004) that, while the definition may work in the case of the 3rd person (especially when used anaphorically), it is somewhat problematic to think of 1st and 2nd person pronouns as standing for a noun phrase. Their referent is often defined deictically, so that one should think of these pronouns as indexes of a speech-act participant rather than a phrasal constituent.

Personal pronouns constitute an entire constituent of the clause, i.e. they do not modify other elements and they admit no modification of their own. The following examples - taken from a story featuring a personified monkey - show that personal pronouns are most vehemently denied an attributive function:

(52)   a.   *yo vay*   *nə*   *kazəmẽt də mĩ́*   *irmã.*
          1s   go.NPST LOC wedding of 1s.POSS sister
          'I am going to my sister's wedding.'

b. * *yo makak vay nə kazəmɛ̃t də mĩ̃ irmã.*
1s monkey go.NPST LOC wedding of 1s.POSS sister

This is a characteristic that personal pronouns share with nouns. However, personal pronouns differ from nouns in fundamental ways, including the fact that they belong to a closed class rather than an open one and that their actual reference must be anchored in discourse/context. Furthermore, personal pronouns cannot be modified; (53) exemplifies the non-acceptability of an adjective, a numeral[20] and a demonstrative as modifiers of a personal pronoun.

(53)   a.   *elz a   sa-i.*
            3p IRR.NPST leave-INF
            'They will come out.'

       b.   * *gran elz a   sa-i.*
            big   3p IRR.NPST leave-INF

       c.   * *doy elz a   sa-i.*
            two 3p IRR.NPST leave-INF

       d.   * *ikəl   elz a   sa-i.*
            DEM$_d$ 3p IRR.NPST leave-INF

In other respects, personal pronouns behave as NPs: case marking can attach both to an NP and a pronoun, as shown in (54) for the dative-marking preposition *pə/a* (sections 7.4.1 explains the distribution of the two forms):

(54)   a.   *Vera kõt-o   ũ istər pə Leslie.*
            V.   tell-PST one story DAT L.
            'Vera told a story to Leslie.'

       b.   *Vera kõt-o   ũ istər a   el.*
            V.   tell-PST one story DAT 3s
            'Vera told a story to him.'

The only person/number combination for which there is a distinct oblique form is first person singular, as indicated in Table 6.2 below; example (55) exemplifies how the oblique

---

[20] In (53c), the numeral occurs in what would be its prototypical function as a nominal modifier, and as such is ungrammatical in DIP. Co-occurrence of the two elements such as the one given below, in which the numeral occurs after the pronoun, are in fact instances of apposition and therefore constitute no counter-evidence to the claim:

(1)   *ɔn   foy   dəpəy nɔs, nɔs doy?*
      where go.PST then 1p   1p two
      'Where did we go then, the two of us?'

form *mĩ̃* is used in case-marked contexts instead of the direct form *yo*:

(55) *el kɛr   kum-e   a   mĩ̃/*yo nə   istrad.*
      3s want.NPST eat-INF DAT 1s   LOC road
      'He wants to eat me by the road.'

The DIP personal pronoun paradigm can now be given in its entirety, in Table 6.2. This paradigm is best described as consisting of distinct, largely suppletive, forms for three persons and two number categories, a direct-oblique distinction affecting the first person singular, and marginally also a two gender distinction on the third person singular pronoun only.

	Singular	Plural
1st person	yo, mĩ̃[OBL]	nɔs
2nd person	use	usez
3rd person	el[m.] , ɛl[f.]	e(l)z

Table 6.2: DIP personal pronoun paradigm

Despite the gender and number distinctions made for third person pronouns, *el* can be used by default whether the referent is masculine or feminine, singular or plural. Therefore, while *ɛl* is clearly feminine and *elz* is umistakably plural, *el* is unspecified for those categories. It is interesting to notice that the forms of the DIP personal pronoun contain remnants of the Portuguese plural marker, the suffix *-s*, (in *usez* and *elz*), but the fact that this suffix is not productive at all speaks against such morphological boundary in DIP.

6.6.2 Demonstrative pronouns

In common with personal pronouns, the reference of demonstrative pronouns is usually either deictic or anaphoric. The forms of the demonstrative pronouns are the same as those of the adnominal demonstrative (see section 6.5.4), viz. *es* 'this' (proximal) and *ikəl* 'that' (distal) and *ot(r)* 'other' (absentive). Like personal pronouns, demonstrative pronouns often have deictic and/or anaphoric reference and they are common in copula constructions. The sentences in (56) exemplify the latter construction:

(56) a. *ikəl   ɛ   gran boɖ.*
          DEM$_d$ COP$_i$.NPST big   dumb
          'That [one] is a big dumbo.'

      b. *ot   nã̃   ɛ   bõ̃.*
          other NEG$_{cl}$ COP$_i$.NPST good

'A different one [Lit. 'another one'] isn't good.'

6.6.3 Interrogative Pro-forms

Interrogative pro-forms stand for a number of elements at clause- and phrase-level in questions. Given the heterogeneity of such elements, some must be classified as pronouns while others are pro-adverbs and pro-determiners.[21]

The choice of interrogative pro-form reflects the expected epistemic domain of the element requested: the pro-adverbs *kom* 'how/why' and *pərki/purki* 'why' are used for manner and reason adverbials, *kwɔn* 'when' for time adverbials and *ɔn* 'where' for locative adverbials, while interrogative pronouns *(u)ki* 'what' and *kẽ* 'who' typically stand for non-human/inanimate arguments and human/animate arguments respectively. Inasmuch as they sometimes appear in a non-attributive context, the question words *kwɔl* 'which' and *kwõt* 'how much/many' can be seen as interrogative pro-forms demanding specification and quantification respectively. However, they function more often than not as NP-operators, and as such they are fully discussed in section 6.5.5.

The following sentences exemplify the use of *kwɔn* 'when' (57a) and *kẽ* 'who'(57b) in interrogative sentences:

(57) a. *es   kwɔn kõpr-o?*
         DEM$_p$ when buy-PST
         'When (did you) buy that?'

      b. *ɔy,   pə   Bablu kẽ   vay   cham-a?*
         INTJ DAT B.   who go.NPST call-INF
         'Hey, who is going to call Bablu?'

The same series of pro-forms (*kẽ* 'who', *kom* 'how/why', *kwɔl* 'which', *kwɔn* 'when', *kwõt* 'how much/many', *ɔn* 'where', *pərki* 'why' and *uki* 'what') also intervene in complement clauses. Such constructs usually convey non-specific reference, although that need not be the case:

(58) *yo uki fal-a   use ti   ki   faz-e.*
     1s what say-INF 2   AUX.NPST CMP do-INF
     'You must do whatever I say.'

Kwɔn and *pərki* are also subordinators (see section 6.8.2). Furthermore, there is partial formal overlap between DIP's interrogative series and its relative pronouns (see 6.6.4).

---

[21]One interesting aspect of the DIP interrogative pro-form paradigm is that it makes little use of bimorphemic question words; this is striking if we consider that both Daman IP and Korlai IP do (Clements and Koontz-Garboden 2002:210), as well as the extinct Norteiro varieties (Dalgado 1906); the same is true of other Asian creoles, such as Papia Kristang (Baxter 1983).

6.6.4 Relative pronouns

In a relative clause, the relativised element is almost always expressed by the relative particle *ki*, as in (59):

(59) *es   ɔfis? dəpəy d-ikəl   ki   te   ũ   kaz?*
      DEM$_p$ office after of-DEM$_d$ REL EXS.NPST one house
      '[Do you see] this office? [Do you see] a house which is after that?'

Apart from some restrictions, such as the impossibility to co-occur with prepositions (see 7.10.3), *ki* is used regardless of the function of the relativised element in the relative clause; there is, in addition, a complete series of relative pronouns with different forms reflecting their role within the relative clause. The use of most of these relative pronouns is somewhat hesitant.

The forms of most relative pronouns are similar to the interrogative pro-forms (see 6.6.3), except that the form *uki* mentioned in connection with the inanimate/non-human interrogative pronoun is excluded from a relativising function and always surfaces as its competing form *ki*. The list of relative pronouns includes *kẽ* 'who' for human/animate arguments, *ki* for non-human/inanimate arguments, NP-internal *kwɔl* 'which, who', *kom* or *kufɔr* 'how' (usually relativising *maner* 'manner') if the NP has a manner adverbial function in the relative clause, *kwɔn* 'when' (usually relativising time units such as *di* 'day' or *ɔr* 'hour') if it refers to time and *ɔn* 'where' if the NP participates of a locative phrase. The distributional equivalence of the relative particle relative pronouns is exemplified for *kom/kufɔr* in (60) and *kwɔn*:

(60) a. *asĩ   də maner   ki/kom/kufɔr nɔs fal   ẽ   kaz.*
          thus of manner REL   1p speak.NPST in house
          'Thus, the way that we speak at home.'

      b. *kwɔn ε   last time ki/kwɔn nɔs   foy damãw?*
          when COP$_i$.NPST last   time   REL 1p go.PST Daman
          'When was the last time that we went to Daman?'

See section 7.10.3 for further information on the syntactic charcteristics of relative clauses.

6.6.5 Indefinite pro-forms

Indefinite pro-forms convey a number of meanings associated with indefiniteness, ranging from unknown reference (e.g. *someone*), indication of free choice (e.g. *anywhere*) or non-existence (e.g. *nobody*). Haspelmath (1997) notices that, in several languages, these diverse meanings are distributed over various parallel series with important formal similarities, and also that indefinite pro-forms tend to indicate the epistemic domain of the replaced elements

in much the same way as interrogative pro-forms[22] (from which indefinite pro-forms often derive, cross-linguistically).

DIP makes use of two series, viz. the *aŋ-* series expressing both lack of knowledge and free choice, and what can be termed the negative *n-* series.

In the *aŋ-* series, the only real pro-form is the pronoun *aŋe* 'somebody', indexing specific animate (usually human) referents. All other indefinite reference is expressed through analytical constructions consisting of the existential quantifier *aŋũ* 'some, any' (optionally also *kwɔlki* or *sɛrt* 'some, any') plus a generic nominal reference to the desired epistemic domain:

(61) a. *aŋe*   *fez*   *fon.*
          somebody make.PST phone
          'Somebody called.'

      b. *aŋũ/kwɔlki/sɛrt jẽt   pəd   a br-i   kok   sẽ   fak.*
          some   people can.NPST open-INF coconut without knife.
          'Some people can open coconuts without a knife.'

      c. *kurəsãw də makak   kwɔlki di   ad   sa-i   fɔr?*
          heart   of monkey some day IRR.NPST leave-INF out
          'Would a monkey's heart ever come out?'

The *n-* series, that of the negative indefinite pro-forms, contains suppletive simple forms to replace animate (usually human) referents (*nĩge* 'nobody'), non-human referents (*nad* 'nothing') and also for the epistemic domain of time (*nuk* 'never'). For all other domains, DIP makes use of analytical constructions with the NP negator *niŋũ* 'no [X]':

(62) a. *də nĩge   yo nə   tə   fik-a   med.*
          of nobody 1s NEG$_{cl}$ IPFV.NPST become-INF fear
          'I am not scared of anyone.'

      b. *nad   nãw.*
          nothing NEG$_{cl}$
          'Nothing.'

      c. *yo nã   foy   niŋũ part.*
          1s NEG$_{cl}$ go.PST no   place
          'I didn't go anywhere.'

---

[22] Karcevski (1969b,a) (see also Evans 2000, Mackenzie in press) introduces the category of *ignoratives* to account for words which combine the functions of interrogative and indefinite pro-forms, and possibly also relativisation, negation or anaphoric reference.

Excluding the suppletive forms, it therefore emerges that, from a formal point of view, the DIP indefinite pro-forms are largely generic-noun-based (cf. Haspelmath 2005a).

6.7 Prepositions

DIP is predominantly a prepositional language, as all adpositions (with the possible exception of one, see below) occur prenominally. Prepositions - some of which are highly macrofunctional (see 8.1.3 for a definition of the concept) - may signal syntactic/semantic roles, and establish the relationship between two or more participants. Their use and functions are described in more detail in the following chapter (in particular section 7.1)

Simple prepositions include the following: *a* 'to, at [temporal]', *ata/(a)tɛ* 'until', *də* 'of', *ẽ* 'in' (rarely used), *ku/ko*[23] 'with', *sẽ*[24] 'without', exemplified in (61b) above, *kom* 'like', *kufɔr* 'the same way as', *nə*[25]*/nu* 'in' and *pə/pu* 'to'.

These prepositions directly precede the NP or pronoun they govern, without any further intervening elements, as exemplified for atɛ in (63):

(63) *el fuji-w   i   foy   atɛ   ikəl   line.*
      3s flee-PST and go.PST until DEM$_d$ line
      'He ran and went until the (finishing) line.'

In addition, there is also a series of complex prepositions in which *də* combines with a subset of lexical items (semantically close to what in other languages are called *relational nouns*, such as 'inside', '(in) front' or 'top'). The preposition *də* in these compounds is often extremely reduced from a phonetic point of view and may be left out altogether - if the trend generalises, one will begin to interpret the previously adverbial/nominal forms as grammatical elements, but that does not seem to be the case as yet. Some examples are given in (64):

(64) a. *yo gɔs   kõt-a   istɔr jŭt   (də) Conchita.*
          1s like.NPST tell-INF story together of   C.
          'I like to chat [Lit. 'tell stories'] to Conchita.'

      b. *use ẽtr-a   dẽt   d-es   daba.*
          2   enter-INF inside of-DEM$_p$ tin
          'You'll get into this tin.'

Complex prepositions include *ãt də* 'before', *bayx də* 'under', *dəpəy də* 'after', *dẽt də* 'inside', *durãt də* 'during', *fɔr də* 'outside', *frẽt də* 'in front of', *jŭt də* 'COMITATIVE', *(pu)kawz*

---

[23]Not to be mistaken for the NP connector, see 6.8.

[24]Not to be mistaken for the homophonous cardinal numeral, see 8.4.1

[25]Not to be mistaken for the pragmatic tag/requestative or negator, see 7.7 and 7.6.2.2.

də 'because of', *lad də* 'next to', *lõj də* 'far from', *pɛrt də* 'close to', *(i)sim də* 'on top of' and *(a)tras də* 'behind'.

The prepositions *nə*, *də* and, to a lesser extent, *pə* often contract when they precede a vowel-initial[26] word such as a demonstrative: *nə/də/pə + es* results in *nes* 'in this [...]', *des* 'of/from this [...]' and *pes* 'for this [...]'.

One adposition sometimes contradicts the canonical PREP + NP word order: *jŭt də* 'COMITATIVE' sometimes occurs in a *(də)* + NP + *jŭt* construction:

(65) *mĩ   jŭt   nã   te   muyt diŋer nã   te.*
1s.OBL together NEG*cl* EXS.NPST much money NEG*cl* EXS.NPST
'I don't have much money.' [lit. with me/next to me there isn't much money.]

This sort of inversion is reminiscent of similar inversions involving the pre-nominal occurrence of a possessive *də*-PP (in which case the preferred composition is simply a bare [Oblique] personal pronoun); it is likely that the inversion of the components of the complex adposition *jŭt də* occurs by analogy with what has been described for the possessive PP. Other complex prepositions do not show the same flexibility in terms of word order.

6.8 Connectors

DIP connectors fall under two categories: *coordinators*, which connect elements (phrases or clauses) with a similar syntactic status, and *subordinators*, which introduce subordinate clauses.

6.8.1 Coordinators

The most common coordinator in DIP is *i* 'and', which can conjoin NPs, predicates, clauses (66a) or any other units. The coordinator *ku/ko* 'and' (equivalent to the comitative/instrumental preposition) is reserved for NP conjunction (66b). The adversative coordinator is *may* 'but' (66c)and the disjunctive coordinator is *o* 'or' (66d):

(66) a. *nɔs põy   pimẽt, dəpəy põy   pok grãw, ayl, jĩjiv, tud əkəl*
1p   put.NPST pepper then   put.NPST little gram garlic ginger all DEM*d*
i   mɔy.
and grind.NPST
'We put pepper, then a bit of gram, garlic, ginger, all that, and (we) grind (it).'

b. *nə   Go   yo te   bastãt cousin ku auntie.*
LOC Goa 1s have.NPST many cousin and auntie

---

[26]This characteristic of *də* is at the root of a theoretical problem concerning the categorisation of some possessives (*duse*, *del* and *dɛl*) as independent grammatical forms or prepositional phrases (see 6.5.4).

'In Goa, I have many cousins and aunties.'

c. *nɔs nə   pɔd   kum-e   muyt may nɔs kɔm.*
   1p NEG$_{cl}$ can.NPST eat-INF much but 1p eat.NPST.
   'We can't eat a lot but we do.'

d. *use kɛr   brãk kamis o pret?*
   2   want.NPST white shirt or black
   'Do you want the black shirt or the white one?'

6.8.2 Subordinators

Common DIP subordinators include *kwɔn* 'when', introducing time clauses, *pures* 'therefore', which indicates that the subordinate clause introduces a logical consequence of the main clause *pərki* 'because', introducing clauses of reason, *si* 'if', indicating a conditional clause, and *pu* (occasionally *pə* 'to'), which marks the embedded clause in purposive constructions. The latter, however, also doubles as a complementiser introducing non-finite clauses. The double use of *pu* is exemplified in (67):

(67) a. *el t-iŋ   vay   nə   ũ   jungle pu   traz-e   koys.*
      3s IPFV-PST go.INF LOC one jungle PURP bring-INF thing
      'He went to a jungle to bring [some] things.'

    b. *fil   də Gilbert pidi-w   a   mĩ   pə   ajud-a   a   ɛl.*
      child of G.   ask-PST DAT 1s.OBL PURP help-INF DAT 3f
      'Gilbert's daughter asked me to help her.'

Complementisers are specific types of subordinators which mark complement clauses rather than adjuncts (see 7.10.1). Another DIP complementiser is *ki*, which introduces object complement clauses (68a) and even utterance complements (68b):

(68) a. *yo sab-iŋ   ki   el tə   fal-a   Liza mem.*
      1s know-PST CMP 3s IPFV.NPST say-INF L.   EMPH
      'I knew that he would say Liza only.'

    b. *i   dəpəy lion fal-o   ki   muyt ubrigad.*
      and after lion say-PST CMP much thank.you
      'And then the lion said: 'Thank you very much'.'

Subordination is a complex matter in DIP which, apart from the elements described here, makes use of the full series of the interrogative pro-form series as well. For a complete description of subordination, see 7.10.

6.9 Other particles

In addition to the uninflected grammatical words described earlier (connectors, prepositions, certain pro-forms), DIP makes use of a number of other invariant particles, with very disparate functions. I will propose a distinction between particles which participate in the structure of a clause or phrase (grammatical particles) and those which normally constitute a complete utterance in themselves (interjections).

6.9.1 Grammatical particles

The invariable particle *a(d)* is the non-Past Irrealis marker, which often indicates Future tense (69a). *Vidi* is its Past counterpart, particularly active to express counterfactual mood (69b) - see 7.1.2.2.

(69) a. *nɔs doy a   faz-e   race.*
          1p   two IRR.NPST make-INF race
          'We will make a race.'

      b. battery *aki   si te,   chalu   vidi   fik-a.*
          battery here if EXS.INF functioning IRR.PST become-INF
          'If there were a battery here, (it) would be on.'

The use of the relative particle *ki* 'that' has already been treated in 6.6.4. It is normal to enforce a terminological division between a relative particle and relative pro-forms in that the former is invariable whatever the role of the relativised element in the relative clause, whereas relative pro-forms reflect this in some way.

Mem is an emphatic particle which attaches to several kinds of clausal constituents (predicates, NPs, adverbials) as well as phrase-level elements (numerals, possessives, adjectives, demonstratives). The use of *mem* is fully spelt out in section 7.8.5.

The requestative particle *nə* is highly macrofunctional, participating in the formation of polar questions (70a), imperatives (70b), and to demand general acknowledgment of the clarity of an utterance or constituent (70c):

(70) a. *liza õt   nã   t-iŋ   med nə   pə   uncle.*
          L.   yesterday NEG$_{cl}$ have-PST fear REQ DAT uncle
          'Yesterday Lisa was not afraid of uncle, right?'

      b. *fal   nə   nə   uvid.*
          speak.NPST REQ LOC ear
          'Do speak in my ear.'

c.   *ali*   *yo t-iŋ*   *gi-a*   *saykəl nə*   *yo vi-w*   *pə*   *use.*
        there 1s IPFV-PST ride-INF bicycle REQ 1s see-PST DAT 2
        'I was riding the bicycle there, you know, [and] I saw you.'

For a full account of its functions, see sections 7.6.2.2 and 7.6.3.1.

The clausal negator *nã/nə* 'not' modifies the predication in order to establish negative reference - see examples (62c), and (70a) above, also section 7.7. Its emphatic counterpart, *nãw*, is in most instances best treated as an interjection (see 6.9.2).

The hortative particle *bam* 'let's' normally precedes a non-finite verb form to indicate a strong suggestion, decision or encouragement. For a complete description of the applications of *bam*, see section 7.6.3.3. This particle is sometimes used as an interjection, although this is not its most common function.

The restrictive particle *sɔ* 'only' and the additive particle *time/timẽ* 'also' have opposite functions, exemplified in (71):

(71)   a.   *el tə*   *beb-e*   *sɔ*   *leyt.*
            3s IPFV.NPST drink-INF only milk
            'He drinks nothing but milk.'

       b.   *ε*   *Jacob, a*   *mĩ*   *time vẽy*   *i*   *sigər-a.*
            INTJ J.   DAT 1s.OBL also come.NPST and catch-INF
            'Hey, Jacob, come and catch me too.'

The use of the restrictive and the additive particle is described in detail in 7.8.5.

6.9.2   Interjections

The emphatic negator *nãw* 'no' is an interjection which appears most commonly in syntactic isolation - a corresponding non-emphatic form is treated in 6.9.1 as a negative polarity particle. Likewise, the affirmative polarity particle *sĩ* 'yes'. Both of these suffice to answer a polar question (see 7.6.2.1).

DIP also makes use of a number of areal interjections. They have different pragmatic functions, such as that of conveying surprise or irritation (e.g. *are*); exasperation (e.g. *ufo*); to express that one is nervous (e.g. *baba*); or simply to attract attention - e.g. ε and *ɔy* (see examples (57b) and (71b)).

Chapter 7

Syntax

This chapter describes the most salient features of DIP syntax as well as the interaction between syntax and the expression of specific pragmatic values. I will begin by describing the organisation of elements at the phrasal level; the predicate phrase (including both verbal and non-verbal predication) is described in section 7.1, the noun phrase in 7.2, adjective phrases and adverbial phrases are treated together in 7.3, while the structure of prepositional phrases is described in 7.4. The description goes on to the language's strategies of argument alignment, in section 7.5, before exploring the basic structure of simple clauses, in 7.6. The syntax of negation is described in 7.7, while section 7.8 deals with the interaction between pragmatic status and syntactic (as well as prosodic) structure. Complex sentences are explored in sections 7.9, on coordination, and 7.10, concerning the various types of subordinate clauses. Finally, section 7.11 describes the syntax of comparative and similative constructions.

7.1 The predicate phrase

The scope of this section accommodates both verbal predication - i.e., the structure of Verb Phrases (VPs) - as well as non-verbal predication - in the form of copular constructions.

The verb phrase minimally consists of a verbal form (be it a lexical verb or a copula). The main verb determines to a large extent the argument structure of the clause, and it can also be modified by adverbials - the latter are treated in section 7.3.

Tense, aspect and mood marking is at times morphological and often periphrastic. The locus of free TAM operators is to the left of the main verb, which is occupied by auxiliaries (particles or grammatical verbs). With all auxiliaries other than *tə/tiŋ*, the infinitive form of the main verb is used; *tə/tiŋ* may combine with a participial form of the main verb - although this construction is rather marginal -, but in fact this particular verbal element transcends the category of auxiliaries because of its close connection with the macrofunctional verb *te*, used as a copula as well as in existential, locative and possessive constructions (see 7.1.3). The implications of this fact for the interpretation of participle constructions is explored in section 7.1.5.

To the left of any potential non-morphological TAM markers, the VP still accommodates polarity items with clausal scope (see 7.7). No clausal constituent can occur in

between these elements and the main verb, which attests to their structural unity. Phasal adverbs are usually realised in a similar position, although they show some degree of placement flexibility (see 7.6.4).

DIP verbs have two finite forms, which do not participate in auxiliary constructions. This section therefore treats the occurrences of finite verbs and auxiliary constructions separately. The distribution of the different types of constructions reveals certain semantic subtleties concerning the type of predication (State of Affairs, in the terminology of Dik 1997) involved. It is therefore important to start by introducing some of the relevant concepts.

The notions of *stativity* and *telicity* are particularly important to understand the use of some of the most common DIP auxiliaries. Predication (rather than verbs) can be classified as *stative* or *dynamic* depending on whether it effects a change in state of affairs. According to this criterion, the predication associated with a verb such as 'to want' is typically stative whereas that conveyed by 'to bake' is usually dynamic.

Dynamic predications can be further classified according to *telicity - boundedness* in the terminology of Östen Dahl (1985) -, i.e., whether or not the action is perceived to have a clear endpoint. Telicity is not necessarily a semantic property of the verb but rather of the predication. Consider the following sentences:

(72)   a.   Eva baked bread for a living.

          b.   Eva baked a loaf of bread.

In (72a), baking bread is construed as a habitual activity, i.e. one without a clearly defined endpoint, and as such constitutes an *atelic* predication. Inversely, the activity of baking as expressed in (72b) can be assigned a clear moment of completion, and the predication is accordingly considered *telic*. Stative verbs and telic actions share the property that their internal temporal structure is of minor importance: undefined for the former, irrelevant for the latter. As a result, stative and telic predicates often reject imperfective marking. Recognising this fact will help us understand the distribution of finite verbs, on the one hand, and the *tə/tiŋ* pair of auxiliaries in particular.

Having set out the aims of this section, as well as some relevant concepts, I will begin by describing the occurrence of finite verb forms and their Tense and Aspect characteristics, in 7.1.1. The following subsection, 7.1.2, introduces the grammatical elements which modify the nonfinite form of the main verb to express Tense and Aspect. Section 7.1.3 describes non-verbal predication and its associated constructions, while 7.1.4 deals with conjunct verbs. I will then address the question of the status of participial constructions in section 7.1.5 - including a note on the irregular occurrence of gerunds.

7.1.1   Finite verbs

As explained in section 6.2, we can recognise two distinct finite forms for DIP verbs, whether suffixed or suppletive, which code a Past/non-Past tense distinction. The follow-

ing examples contain several finite verb forms, both Past and non-Past; for a lexical pair, compare the form *kõp* in (73a) with *kõpro* in (73b):

(73) a. *tud   di   fik   ukupad   pərke   nɔs kõp   koyz   də sɛman nə.*
whole day stay.NPST occupied because 1p buy.NPST thing of week REQ
'We are busy the whole day day because we buy everything for the week, you see?'

b. *dəpəy* afternoon *el   foy   tros   galiŋ,   kõpr-o   kok,   fez*
then afternoon 3s go.PST bring.PST chicken buy.PST coconut make.PST
chatni.
chutney
'Then, in the afternoon, he went and brought chicken, he bought coconut and made chutney.'

The option for the label *non-Past* rather than *Present* derives from the fact that most of the occurrences of these forms refer to *Habitual* actions, as in (73a), or *Generic* predications, as in (74):

(74) *nə   damã̃w nã̃   fal   biting.*
LOC Daman NEG$_{cl}$ say.NPST biting
'In Daman [they] don't say "biting".'

In addition to tense, finite verbs in DIP also contribute aspectual information, defined in contradistinction with periphrastic verb groups containing the auxiliaries *te/tiŋ*. The latter are indicators of Imperfective aspect (cf. section 7.1.2.1). In the Past tense, at least, finite verbs express Perfectivity, as shown in (73b) and the following contrasting pair of sentences:

(75) a. peacock *uv-o.*
peacock fly-PST
'The peacock flew.'

b. peacock *t-iŋ   uv-a.*
peacock IPFV-PST fly-INF
'The peacock was flying.'

Notice that the verb in (75) is dynamic. Stative verbs, as mentioned earlier, normally reject imperfective aspect and therefore occur in their finite form in both non-Past (76a) and Past (76b):

(76) a. *yo gɔs   mãg.*
          1s like.NPST mango
          'I like mangoes.'

      b. *yo gost-o   muyt.*
          1s like-PST much
          'I liked [it] very much.'

The fact that statives seemingly reject Imperfective marking in non-Past as well brings us to the issue of whether or not non-Past finite verb forms can be considered to imply Perfective aspect. This may be so, but the issue merits further commentary. Whereas it is true that, in non-Past, the periphrastic *tə* + INF construction does draw attention to the ongoing nature of the event, the punctual nature of Present reference makes it difficult to conceive of a completed Present action - and, as we shall see, Future reference has its own separate instruments. The non-Past form of verbs is employed for stative and atelic predication (cf. example (73a)), not because the actions are to be perceived as rounded-off but because their internal temporal structure is irrelevant. This intuition is in tune with common definitions of the concept of Perfectivity, according to which Perfective is characterised by a holistic view of the predication rather than its completeness. Comrie offers the following summary:

> perfectivity indicates the view of a situation as a single whole, without distinction of the various separate phases that make up that situation, while the imperfective pays essential attention to the internal structure of the situation. (Comrie 1976:16)

The Perfective/Imperfective opposition is an essential one in DIP, which interestingly marks perfectivity morphologically and imperfectivity periphrastically. The semantic boundaries between the two categories are not entirely impermeable, as certain aspectual categories cross them more or less freely. Even though Habituals are more often than not marked with the Imperfective auxiliary[1] - see (77b) as well as (86) below -, notice how a finite verb may have a similar reading, in (77a); no comparable example has been recorded with a finite Past tense verb form.

(77) a. *Liza sẽp   bɛb   cha ãt   də vay   durm-i.*
          L.   always drink.NPST tea before of go.INF sleep-INF
          'Liza always drinks tea before going to sleep.'

      b. *el tə   beb-e   sɔ leyt nə.*
          3s IPFV.NPST drink-INF only milk REQ
          'He drinks nothing but milk, you see.'

---

[1]Similarly in Daman IP, see Clements and Koontz-Garboden (2002:216ff).

Verb forms morphologically marked for imperfectivity are rare and highly acrolectal. The corresponding suffix -*v* is equivalent to the standard Portuguese imperfective marker -*va*-. Some examples taken from recordings of unconstrained speech are given below.

(78) a. *nə   natal   da-v   aroz, asuk, cha, mãteg, azet, kar.*
          LOC Christmas give-IPFV rice sugar tea butter oil   meat
          'For Christmas (they) would give (us) rice, sugar, tea, butter, oil, meat.'

      b. *dumĩg pasad kwɔn ista-v   aki ɛ-r   mĩ   an.*
          sunday past when EXS-IPFV here COP$_s$-PST 1s.POSS year
          'Last sunday, when you were here, it was my birthday.'

(78a) was produced by an elderly lady whose speech provides further indications of high exposure to standard Portuguese. Example (78b), on the other hand, was spoken by a middle-aged lady.

7.1.2   Auxiliary constructions

Auxiliary constructions are characterised by the combination of the infinitive form of the main verb with a preposed auxiliary, which contributes the necessary tense, aspectual and/or modal information. Auxiliaries can in most cases be organised in pairs of distinct forms for Past and non-Past reference.

Many auxiliaries are particles rather than verbs, although often derived from verbs. Others, however, are transparently related to lexical verbs. It is often challenging to decide whether a V_{fin} + V_{inf} sequence refers to an instance of auxiliary modification (i.e., whether the V_{fin} is purely grammatical) or rather of a lexical verb governing an nonfinite complement clause with no overt complementiser (i.e., a *catenative* verb construction, see 7.10.1). In other words, one needs to decide whether a combination of two verbal elements constitutes a monoclausal VP (Anderson 2006) or whether it indicates a combination of two clauses, and in the latter case whether or not one is subordinate to the other.

It is possible to place the co-occurrence of two verbal forms without resort to any overt conjoining element on a scale of *clausal integration* ranging from complete integration (auxiliary constructions, monoclausal constructions in which the auxiliary verb has purely grammatical meaning) to complete independence (coordinated clauses, in which the verbs head clauses of equal pragmatic status). The status of the verb forms is different in different positions of the scale. On the auxiliary end of the scale, one of the verb forms is entirely grammatical, i.e. it carries no semantic meaning; on the independent clause extreme of the scale, on the other hand, both verb forms are equally semantically charged.

From a morphological point of view, the verbs in coordinated clauses with no overt conjoining element - also interpretable as serial verb constructions - all admit inflection; these constructions therefore constitute a very particular phenomenon (discussed in 7.9.1) which is disregarded in the present discussion. Let us therefore concentrate on constructions in which one of the verbs is inflectionally or externally TAM marked while the other

is nonfinite. We now have, in opposite ends of the scale, auxiliary constructions and non-finite complement clause constructions. Some instances are easy to classify as auxiliary constructions. Consider in (79) the case of *tiŋ*, a Past Imperfective auxiliary roughly corresponding to the Past form of the transitive verb of possession *te* 'to have', the intransitive existential verb *te* 'to exist' (see 7.1.3.3) and stage-level copula *te* (see 7.1.3):

(79)   ũ *lion* t-iŋ kur-e.
       one lion IPFV-PST run-INF
       'A lion was running.'

Tiŋ (the non-Past corresponding form of which is *tə*) is treated as an auxiliary because in that capacity (and as a copula) it is basically devoid of semantic meaning, instead carrying grammatical (aspectual) information only, and it furthermore shows evidence of phonological bleaching when compared to its lexical counterpart.

The verb *pude* 'can/could' does not admit any full complement clause or any nominal complement, and it is therefore treated as a modal auxiliary. Its occurrence is exemplified in (80) - see also 7.1.2.3 for further information:

(80)   *nɔs nə*   *pɔd*   *kum-e   muyt.*
       1p   NEG*cl* can.NPST eat-INF much
       'We can't eat a lot.'

Other verbs in such constructions can be shown to fall on the catenative end of the scale. It is often the case that unmarked, nonfinite complement clauses only occur when the subject of the two clauses is identical, and a different construction is required when that is not the case. This is exemplified in (81) for the verb *gosta* 'to like':

(81)   a. *yo   gɔs   brĩk-a   kavok ku   arv.*
           who like.NPST play-INF cave   INS tree
           'I like to play in the caves with the trees.'

       b. *pures   yo nã   gɔs   ki   use tum-o.*
           therefore 1s NEG*cl* like.NPST CMP 2   take-PST
           'That's why I'm not pleased [Lit. 'I don't like'] that you had [something to eat before].'

       c. *yo gɔs   panir.*
           1s like.NPST paneer
           'I like paneer.'

The fact that the verb *brĩka* is nonfinite in (81a) is therefore a function of the coreference of the subject in both clauses involved; were that not the case, the complement clause

would require an overt complementiser and would be finite, as shown in (81b). Notice also that the tense characteristics of the main clause and complement clause in (81b) do not match, which testifies to their relative independence from each other. In support of the interpretation of *brĩka kavok ku arv* in (81a) as a complement clause, the monotransitive nature of the verb *gosta* is shown in (81c), with the occurrence of a nominal complement.

The problematic cases lie in between these poles. One such case is that of the verb *dixa* 'let'. This verb differs from catenatives such as *gosta* in that, even if the actor of the two actions is not equivalent, it does not admit the complementiser *ki* and the second verb form is always nonfinite. This is shown in (82):

(82) a. *el nã   dex   fal-a   nɔs nə   ĩglix.*
          1s NEG$_{cl}$ let.NPST speak-INF 2s   LOC English
          'He doesn't let us say [it] in English.'

      b. *dex   a   mĩ   uvir   ikəl   kãtig.*
          let.NPST DAT 1s.OBL listen.INF DEM$_d$ song
          'Let me listen to the song.'

In fact, verbs such as *dixa* 'to let' (as well as *ajuda* 'to help' or *mãda* 'to command', for example) can be seen to involve a beneficiary argument which is coreferential with the agent of a different action, expressed in a nonfinite clause.[2] I will therefore treat them as catenative verb constructions, even if these fall somewhere in the middle of the clausal integration scale - see section 7.10.1 for further information on the use of *dixa*, including its case assignment properties.

Certain verbs have both an auxiliary use and a catenative use, in which case there is a clear-cut distinction between the two. This is the case with the verb *kere*. In its catenative use, *kere* means 'to want', and it admits a complement clause which is nonfinite if its referent coincides with that of the main clause (83a) and finite otherwise (83b); *kere* is also a regular transitive verb, admitting nominal complements (83c):

(83) a. *el kɛr   tum-a   interview.*
          3s want.NPST take-INF interview
          'He wants to interview [us].'

      b. *el kɛr   ki   nɔs kõt   istɔr.*
          3s want.NPST CMP 1p   tell.NPST story
          'He wants us to chat.'

---

[2] Notice that the logical actor of the second verb is unmarked in (82a) - *nɔs* - which is a characteristic of grammatical subjects; on the other hand, in (82b) *a mĩ* is dative-marked, which might suggest it is a beneficiary/object of *dex* rather than the grammatical subject of *kume* (see 7.5). Despite this intepretative duality, we are dealing with a complement clause because, from a semantic point of view, the two verb forms integrate two different propositions.

c. *el kɛri-w   ag.*
    3s want-PST water
    'He wanted water.'

The auxiliary use of *kere* is rather different, indicating deontic modality. An example is given in (84), and further information can be found in 7.1.2.4:

(84) *yo a   ker-e   kõt-a   istɔr jŭt   də Steven.*
      1s IRR.NPST want-INF tell-INF story together of S.
      'I must have a chat with Steven.'

This discussion exemplifies the reasoning behind my treatment of V_{fin} + V_{inf} sequences. All forms (verbal or otherwise) identified as auxiliaries are individually described in the following subsections.

7.1.2.1 Auxiliary *te*

As mentioned in 7.1.1 above, the non-Past form of the auxiliary *te* is *tə*, and *tiŋ* its Past counterpart.[3] The auxiliary forms *tə/tiŋ* are peculiar in that they may combine with any nonfinite verb forms, both infinitives and participles. The latter are somewhat marginal and in a sense blur the boundaries between verbal and non-verbal predication, and for this reason they are separately treated in section 7.1.5. In this section, I will concentrate on the dynamics of *tə/tiŋ* within the VP.

The prototypical aspectual meaning of *tə/tiŋ* is Imperfective (often with a Progressive reading), by contradistinction with the Perfective meaning of finite verb forms. This opposition is clearest in the Past tense, as demonstrated in (85) and also in (75) above:

(85) a. *yo gi-o   saykəl.*
          1s ride-PST bicycle
          'I rode my bicycle.'

      b. *yo t-iŋ   gi-a   saykəl.*
          1s IPFV-PST ride-INF bicycle
          'I was riding my bicycle.'

---

[3]These forms establish a close formal link with the existential/copular/possessive verb *te*, which also has *tiŋ* as their Past form. The notational distinction I enforce here for the non-Past forms (*tə* for the auxiliary, *te* for the existential/copular/possessive verb) reflects the fact that, as an auxiliary, the vowel consistently appears noticeably reduced when compared to the other verbs. Still, this may be purely contextual, since auxiliaries are often not stressed in DIP and, furthermore, phonological bleaching usually accompanies processes of grammaticalisation. I do not make too strong a claim concerning the lexical ridge between the auxiliary and the existential/copular/possessive verb, suffice it to recognise that they enter different syntactic templates for different functions, and the realisation of the non-Past form reflects this.

Both progressive actions (86a,b) and habitual predications (86c,d) are commonly marked with the Imperfective auxiliary:

(86) a. *el tə fal-a mĩtir.*
3s IPFV.NPST speak-INF lie
'He is telling lies.'

b. *nə saykəl uncle t-iŋ vay.*
LOC bicycle uncle IPFV-PST go.INF
'Uncle was going by bicycle.'

c. *Conchita ε mɛdroz pərki əl tə fik-a də tud*
C. COP$_i$.NPST fearful because 3s IPFV.NPST become-INF of all
koyz med.
thing fear
'Conchita is very easily frightened because she is scared of everything.'

d. *yo tudi t-iŋ ir pasi-a.*
1s every.day IPFV-PST go.INF tour-INF
'I would go on tours everyday.'

Considering how diffuse a variety DIP is, it is no surprise to find counter-examples to the prototypical aspectual meaning identified for *tə/tiŋ*. In some cases, the action marked by the Past auxiliary *tiŋ* is not necessarily imperfective, as in (87):

(87) *oj də maŋã yo ku Fabian t-iŋ gi-a saykəl.*
today of morning 1s and F. IPFV-PST ride-INF bicycle
'This morning Fabian and me rode our bicycles.'

In certain instances, then, this analytical construction is interchangeable with the Past-inflected finite form of the main verb. In one recording session, several children were simultaneously asked *use õt uki fez?* 'what did you do yesterday?', and both sentences in (88) were given as answers:

(88) a. *yo t-iŋ brĩk-a saykəl.*
1s IPFV-PST play-INF bicycle
'I played with my bicycle.'

b. *yo brĩk-o saykəl.*
1s play-PST bicycle
'I played with my bicycle.'

The auxiliary *te* may be followed by the particle *ki*, which results in a kind of compound auxiliary *ti ki/tiŋ ki* with a deontic modal meaning of obligation[4]:

(89) a. *ali   də   Purtəgal ti   ki   dar   rɛpos.*
          there from Portugal AUX.NPST CMP give.PST answer
          '[They] have to give an answer, from Portugal.'

      b. *nɔ ti   ki   vay   iskɔl   i   dəpəy nɔs pəd   brĩk-a.*
          1p AUX.NPST CMP go.INF school and then 1p can.NPST play-INF
          'We must go to school, afterwards we can play.'

Sentence (90) exemplifies a comparable occurrence of the Past form of the auxiliary:

(90) *yo t-iŋ   ki   istəd-a.*
      1s AUX-PST CMP study-INF
      'I had to study.'

For a different auxiliary construction with comparable modal semantics, see section 7.1.2.4.

7.1.2.2 Auxiliaries *a(d)/vidi*

Auxiliaries *a(d)* (often *a* and *ad*, but sometimes also *adə* or *ədə*) and *vidi* are Irrealis markers, non-Past and Past respectively. Contrary to *tə* and *tiŋ*, these are particles rather than verbal elements, as there is no synchronic evidence of any inflectional morphology.[5]

Ad is prevalent not only in predictive (91a) and hypothetical Future reference (91b) but also in the apodosis (i.e., the matrix clause) of conditional constructions (91c), prototypical uses of the category of Irrealis (see e.g. Mithun 1995, Romaine 1995b):

(91) a. *yo a   sa-i   pimer ki   də duse.*
          1s IRR.NPST leave-INF earlier COMP of of.2

---

[4] Although this particle *ki* is formally and possibly etymologically equivalent to a complementiser, what follows is not a finite clause but rather an infinitival verb form. This testifies to the integration of this element *ki* into the complex auxiliary, to the extent that it no longer displays its basic complementiser properties.

[5] Although their formal dissimilarity might be interpreted as a case of paradigmatic suppletion, the fact remains that these two elements do not descend from any productive verb in DIP and have no corresponding nonfinite forms. Historically, both particles do derive from the same Portuguese complex modal verb *haver de* 'to be going to/shall', viz. the Present Indicative 3s form *há-de* in the case of *ad*, and from the Past Imperfective 1/3s form *havia de* in case of *vidi*. But the etymological link has been entirely obscured by phonological erosion and a functional speciation which selected these particles but not their original nonfinite counterparts. These elements are therefore treated as invariable particles, even if they can be seen to share an aspectual space.

'I will leave first.' [Lit. 'I will leave earlier than you.']

b. [Why don't you let your little brother ride your bicycle?]

el nə   ad   pud-e   gi-a,   el a   ka-i   vay.
3s NEG$_{cl}$ IRR.NPST can-INF ride-INF 3s IRR.NPST fall-INF go.INF

'He wouldn't be able to ride [it], he would fall down.'

c. *officer si fik-a   sabe-n   nã   ad   gost-a.*
officer if become-INF know-PROG NEG$_{cl}$ IRR.NPST like-INF
'If the officer finds out, he won't like it.'

For further information on the structure of conditionals, including the occurrence of the irrealis markers, see section 7.10.2. In an interrogative sentence, the irrealis marker often expresses doubt/hesitation, in particular in self-directed (92a,b) and counterfactual rhetorical questions (92c):

(92) a. *ŭ   istɔr... kwɔl   a   kõt-a?*
one story   which IRR.NPST tell-INF
'A story... Which one should (I) tell?'

b. *kwõt   an ad   te   fika-d?*
how.many year IRR.NPST IPFV.INF become-PTCP
'How many years has it been?'

c. *kurəsãw də makak   kwɔlki di   adə   sa-i   fɔr?*
heart   of monkey any   day IRR.NPST leave-INF out
'Would a monkey's heart ever come out [of its body]?.'

The particle is also employed to construct dubitative mood, as in (93):

(93) *madriŋ   a   te   sabe-n.*
godmother IRR.NPST IPFV.INF know-PROG
'Godmother should know.'

The Past irrealis marker *vidi* is particularly active in the expression of counterfactual mood. In (94a), the verbal head of the conditional clause codes Past tense, and as such *vidi* can be seen simply as a Past marker of Irrealis. (94b), however, shows that *vidi* occurs in non-Past sentences as well, in which case its counterfactual semantics is the clearest.

(94) a. *elz vidi   faz-e   kaz   si d-elz jŭt   t-iŋ   diŋer.*
3p IRR.PST make-INF house if of-3p together have-PST money
'They would have made a house if they had the money.'

b. battery *aki si te,   chalu   vidi   fik-a.*
battery here if EXS.INF functioning IRR.PST become-INF
'If there were a battery here, (it) would be functioning.'

7.1.2.3 Auxiliary *pude*

This grammatical verb always precedes a nonfinite form of the main verb it modifies. Its Non-Past form *pɔd* 'can' contrasts with the Past form *pudiŋ* 'could'. This auxiliary marks Potential mood in its various manifestations, including the expression of possibility (95a), physical ability (95b) and permission (95c):

(95) a. *es   nã   pɔd   d-a   kum-e pə tud jĕt   də fɔr,*
DEM$_p$ NEG$_{cl}$ can.NPST give-INF eat-INF DAT all people of outside
ɛ   kar.
COP$_i$.NPST expensive
'[One] can't give this for all outside people to eat, it is expensive.'

b. *niŋŭ   peso pɔd   persəb-e   uki   yo fal.*
NEG$_{ct}$ person can.NPST understand-INF what 1s say.NPST
'No one can understand what I say.'

c. *pɔd   kuməs-a.*
can.NPST begin-INF
'[You] may begin.'

Example (95a) shows that the auxiliary remains invariant in negated clauses. According to the rules of negative concord in DIP, described in 7.7 below, one would expect the occurrence of clausal negation in (95b) as well; this is in fact one of the few instances in which constituent negation does not co-occur with clausal negation, but it appears to be an instance of stylistic variation rather than any reaction to the presence of the modal auxiliary.

7.1.2.4 Auxiliary *kere*

The auxiliary/catenative dichotomy of this verb has already been introduced, but it is discussed here in full. As a lexical verb, *kere* is the translational equivalent of the volitive 'to want', both as a transitive verb (96a) and as a catenative verb (96b):

(96) a. *dəpəy el keri-w   ag.*
          then   3s want-PST water
          'Then he wanted water.'

      b. *yo kɛr   fal-a.*
          1s want.NPST speak-INF
          'I want to speak.'

There is however a rather different use of the verb as an auxiliary expressing deontic modality. Modal *kere* occurs in its infinitival form, sometimes bare (97a) but most of the time in combination with an irrealis marker (97b):

(97) a. *use ker-e   ãd-a.*
          2s   want-INF walk-INF
          'You must walk.'

      b. *yo a   ker-e   kõt-a   istɔr jŭt   də Steven.*
          1s IRR.NPST want-INF tell-INF story together of S.
          'I must have a chat with Steven.'

The auxiliary use of *kere* 'must' is equivalent to the *te + ki* construction described in 7.1.2.1. An additional modal meaning is that which could be termed *predictive*. It is not clear whether or not this function requires Irrealis marker, although the few examples encountered do. The following warning, uttered by a young female adult, was directed towards a child who was running in a careless way:

(98) *use a   ker-e   ka-i.*
      2s   IRR.NPST want-INF fall-INF
      'You will (certainly) fall.'

As demonstrated in Khanina (2008), all the meanings associated with this verbal form (both lexical and modal) are in fact normally shared by equivalent forms in various languages across the world.

7.1.2.5 Auxiliary *vay/ir*

As with *kere*, it is also possible to recognise two different uses of the verb *vay/ir*.[6] It is not always easy to tell apart instances of auxiliary *vay/ir* 'be going to' and lexical *vay/ir* 'to go'. The crucial difference is that the former entails no allative/locative semantics, instead

---

[6]Recall that *vay* and *ir* are competing infinitive forms of the same verb. For further information, refer to section 6.2.

signalling both future reference and prediction/intention. As a marker of future reference (and in general), auxiliary *vay/ir* is rather less frequent than *a(d)* (see 7.1.2.2).

The following examples[7] are among the instances of *vay/ir* + V$_{inf}$ for which an allative/locative reading is most clearly unintended:

(99) a. *tudi*   *vay*   *d-a*   *trĩt~trĩt*   *rupi.*
         everyday go.NPST give-INF thirty~thirty rupee
         'Everyday I will give [them] thirty rupees each.'

      b. *dəpəy dis*   *ki*   *faz*   *nə,*   *d-use ɔm* *vay*   *fik-a*
         then say.PST CMP make.NPST REQ of.2   man go.NPST become-INF
         mufin d-es.
         sulky of-DEM$_p$
         'Then [they] said: "Do it, [or] your man will become sulky because of that".'

The verb *vay/ir* is highly prominent in predicate enumeration, in which case the conjoined verbs are often linked by a coordinating conjunction *i* (100a) but not necessarily so (100b):

(100) a. *foy*   *i*   *pasy-o*   *pray,* *kõpr-o*   *grãw i*   *vey*   *kaz.*
          go.PST and take.walk-PST beach buy-PST gram and come.PST home
          '[We] went and took a walk around the beach, bought some gram and came home.'

       b. *dəpəy* afternoon *el foy*   *tros*   *galiŋ,*   *kõpr-o*   *kok,*   *fez*
          then   afternoon 3s go.PST bring.PST chicken buy-PST coconut make.PST
          chatni.
          chutney
          'Then, in the afternoon, he went and brought chicken, he bought coconut and made chutney.'

However, conjoined predicates differ substantially from an auxiliary construction in that both verbs can be finite and/or separately marked for TAM (see section 7.9.1). The problematic situations are those which involve a finite or TAM-marked form of *vay/ir* followed by a nonfinite form of another verb. Unless no allative/locative semantics can be assigned to *vay/ir* in such cases, we are therefore dealing with a catenative verb *vay/ir*.

7.1.3   Non-verbal predication and associated constructions

In DIP, non-verbal predicates are widely attested. In predicative function, non-verbs obligatorily require the support of a semantically poor verbal element. Hengeveld (1992a)

---

[7]See section 7.6.3 for an account of the imperative interpretation of (99b).

makes the point that, in copular constructions, the verbal copula is something of an auxiliary, given that it is the non-verbal part of the predicate to impose selection restrictions. The implication is that copulas are tendentially meaningless, despite the fact that different copulas may be selected by different types of non-verbal predicates.

DIP distinguishes between an *individual-level* copula $\varepsilon/\varepsilon r$ and a *stage-level* copula *te/tiŋ*. As previously pointed out in section 6.5.1, the individual-/stage-level dichotomy reveals a time-stability scale linking prototypical nouns to prototypical adjectives in that predicative nouns usually resist selecting the stage-level copula while prototypical adjectives may or may not occur with its individual-level counterpart - the dichotomy is described in the following subsections, in particular 7.1.3.1 and 7.1.3.2.

As we will see below, the verbal element *te/tiŋ* occurs in various types of constructions, blurring the distinction between a copula and a lexical verb. It has long been established (e.g. Clark 1978, Payne 1997) that in many languages nominal /adjective predicates, existentials, predicate locatives and possessive clauses tend to show formal similarities, in particular with regard to the verbal form employed. Attempting to capture the formal and semantic link between existential, locative and possessive constructions in particular, Clark (1978) calls them *locationals*, arguing that all of these can be construed as indications of (metaphorical) location. These constructions are therefore also described in the present section.

7.1.3.1 Predicate nominals

In this subsection, nominals refer not only to predicate NPs but also pronouns, PPs and subordinate clauses, given that all of these are treated in DIP as full clausal constituents. I exclude adjectives or AdjPs from this batch because they cannot occur outside an NP structure, except in copular constructions. These are described in 7.1.3.2 below. Notice that nominals of this kind can also participate in locative and possessive constructions, for which see 7.1.3.3 and 7.1.3.4 respectively.

Having thus constrained our definition of predicate nominals, the analysis of their distribution reveals that all these select the individual-level copula. The following sentences exemplify the predicate occurrence of NPs selecting the non-Past (101a,d,e,f) and Past (101b,c,g) forms of the individual-level copula:

(101)   a.   *oj* ε *dumĩg.*
            today COP_i.NPST sunday
            'Today is sunday.'

       b.   *ikəl* *di* ε-*r* *mĩ* *an.*
            DEM_d day COP_i-PST 1s.POSS year
            'That was was my birthday.'

       c.   *oj* ε-*r* *last di* *də exam.*
            today COP_i-PST last day of exam

'Today was the last day of exams.'

d. *mĩ* best friend *ε*   *Ashley.*
  1s.POSS best friend COP$_i$.NPST A.
  'My best friend is Ashley.'

e. *Ashley ε*   *mĩ* *amig.*
  A. COP$_i$.NPST 1s.POSS friend
  'Ashley is my friend.'

f. *Jacob ε*   *ũ* *kɔb.*
  J. COP$_i$.NPST one snake
  'Jacob is a snake.'

g. *mĩ* *pay* *ε-r* *pulis.*
  1s.POSS father COP$_i$-PST policeman
  'My father was a policeman.'

The previous examples also show that copular constructions of this type fulfill two rather different functions. In (101a,b,c,d), the predicate nominal is identical to the reference entity; these are known as *equative constructions*. On the other hand, the sentences in (101e,f,g) establish the *proper inclusion* of a given entity into a group of items specified by the nominal predicate. The predicate nominal in proper inclusion copular constructions is often non-specific. The effect of specificity is particularly clear when we contrast (101d) - in which *mĩ best friend* is specific - with (101e) - in which *mĩ amig* is non-specific.

The predicate nominal position may be filled by a PP, as exemplified in (102a,b), or a pronoun, as in (102c)[8].

(102) a. *elz ε*   *də gɔgla.*
    3p COP$_i$.NPST of Goghla
    'They are from Goghla.'

   b. *mĩ* *kaz* *ε*   *pεrt də igrej.*
    1s.POSS house COP$_i$.NPST near of church
    'My house is near the church.'

   c. *mĩ* *kaz* *ε*   *ikəl* *ε.*
    1s.POSS house COP$_i$.NPST DEM$_d$ COP$_i$.NPST
    'My house is that one.'

---

[8]See 7.8.3 for an explanation of the verb doubling observed in (102c).

A possessive pronoun in predicate position must be preceded by a preposition *də* 'of' (see 7.4.3), as in (103):

(103) *es*   *tud koyz* *ε*   *də mĩ.*
       DEM$_p$ PL thing COP$_i$.NPST of 1s.OBL
       'These things are mine.'

A subordinate clause may also occur in the predicate position. The combination in (102), i.e. copula plus a subordinate clause with the subordinator *pu* 'for', transmits intention or expectation.

(104) a. *oj*   *ε*   *pə*   *fik-a*   *dẽt*   *də kwɔt an.*
           today COP$_i$.NPST PURP stay-INF inside of four year
           'Today [we] are supposed to pass to fourth grade.'

       b. *ŭ*   jungly jat *vey*   *i*   *amər-o pε d-ikəl*   *lion, ε-r*   *pu*
           one jungly jat come.PST and tie.PST leg of-DEM$_d$ lion COP$_i$-PST PURP
           kum-e *a*   *el.*
           eat-INF DAT 3s
           'A jungly jat came and tied the leg of the lion, he intended to eat him.'

See also section 7.10.2 for further information on the use of the purposive subordinator *pu*.

7.1.3.2   Predicate adjectives

The clauses containing a predicate adjective are also commonly known as *attributive clauses*. As a result of the loose time-stability of their semantics, predicate adjectives admit both the individual-level copula and the stage-level copula, which makes them the ideal predicates to demonstrate the distinction between the two copulas. An adjective such as *bõ* 'good' is rather versatile in that it may just as easily refer to a state or a more permanent characteristic. Preceded by the individual-level copula as in (105), it is to be interpreted as referring to an intrinsic characteristic of the argument (in this case the speaker) rather than a temporary condition:

(105) *yo ε*   *bõ.*
       1s COP$_i$.NPST good
       'I am [a] good [person].'

Conversely, the selection of the copula *te* in (106) indicates that the predicate refers to a state rather than a time-stable characteristic of the speaker. This is in fact the default reply to the greeting *use kɔm te?* 'How are you?':

(106)   *yo te   bõ.*
       1s COP$_s$.NPST good
       'I am fine.'

Two more examples follow, featuring adjectives other than *bõ*:

(107)   a.   *mĩ   irmãw   ɛ   vɛy.*
             1s.POSS brother COP$_i$.NPST old
             'My brother is old.'

       b.   *es   te   / t-iŋ   vazi.*
             DEM$_p$ COP$_s$.NPST   COP$_s$-PST empty
             'This is/was empty.'

It is AdjPs, rather than adjectives, which occupy the predicate position in these cases. In (108), the adjective *piken* 'small' is modified by an intensifier *tãt* 'so' - see 7.3 for information on the structure of AdjPs:

(108)   *use ɛ   tãt piken.*
       2   COP$_i$.NPST so   small
       'You are so small.'

The forms *te/tiŋ* are interpreted as a stage-level copula when selected by adjectival predicates, but elsewhere they show clear locative semantics. This is explored in 7.1.3.3. It is interesting to notice, however, that the function of the stage-level can be metaphorically linked to location, if one imagines the predicate adjective to establish a bound temporal space within which the veracity of the predication obtains.

7.1.3.3   Existentials and predicate locatives

Existential and predicate locative constructions have a number of formal characteristics in common, to the extent that the latter may be considered an extension of the former. They use the verb *te* and differ only in their expressed valence. Locative constructions are also formally close to possessive clauses, which are described in 7.1.3.4 below.

When used intransitively, *te* has a clear existential reading, as demonstrated in (109):

(109)   a.   *ũ   makak   t-iŋ   i   ũ   crocodile.*
             one monkey EXS-PST and one crocodile
             '[Once upon a time,] there was a monkey and a crocodile.'

b. *nə*   *te*   *nad.*
NEG*cl* EXS.NPST nothing
'No problem.' [Lit. 'There is nothing.']

The sentences above clearly fulfill a presentative function. If this template is extended to include a nominal (110a), adverbial (110b) or phrasal (110c) adjunct whose semantics license a locative reading, then the clause is potentially locative. In such a clause, the locative phrase typically occurs after the verbal form:

(110)   a.   *mĩ*   *irmãw* *te*   / *t-iŋ*   *Go.*
                  1s.POSS brother EXS.NPST   EXS-PST Goa
                  'My brother is/was in Goa.'

         b.   *kwɔn tir-o*   shooting *use ain nã*   *t-iŋ*   *aki.*
              when take-PST shooting 2   yet NEG*cl* EXS.PST here
              'When they did the shooting [of the film], you weren't here yet.'

         c.   *õt*   *yo nã*   *t-iŋ*   (*ẽ*) *kaz.*
              yesterday 1s NEG*cl* EXS-PST in house
              'I wasn't at home yesterday.'

As seen in (110c), a locative phrase need not be marked with a locative preposition. With *te*, bare locative nominals are in fact more common that locative PPs. This is exemplified in (111) below (see also (110a) above):

(111)   *use vi-w*   *n-ikəl*   *di* *kẽ* *t-iŋ*   *mĩ*   *kaz.*
          2   see-PST in-DEM*d* day who EXS-PST 1s.POSS house
         'You saw who was in my house that day.'

Locative *te* therefore commands what could look like an argument, but which is best interpreted as an adjunct. The verb form *te* is not the only one to do so; locative phrases consisting of bare NPs also occur with other lexical locative/locomotional verbs (112a,c,e) such as *fika* 'to stay', *vay/ir* 'to go', *vi* 'to come', *sai* 'to leave' or *ẽtra* 'enter'. These verbs also accept the full range of locative phrases, including spatial adverbs (112d) and locative PPs (112b,f):

(112)   a.   *ĩter* *di* *t-iŋ*   *fik-a*   *kaz.*
                whole day IPFV-PST stay-INF house
               '[She] would stay at home all day.'

b. *yo fik   nə   Diw.*
    1s stay.NPST LOC Diu
    'I live in Diu.'

c. *vay   kaz   nə.*
    go.NPST house REQ
    'Go home.'

d. *Leslie, vay   may tras, may may.*
    L.   go.NPST more back more more
    'Leslie, go further back, further further.'

e. *el ẽtr-o   igrej.*
    3s enter-PST church
    'He entered the church.'

f. *use etr-a   dẽt   d-es   da ba.*
    2   enter-INF inside of-DEM$_p$ tin
    'You'll get into this tin.'

Given that bare locative NPs (112a,c,e) can be replaced with locative PPs, it is wiser to interpret them as special types of locative phrases - licensed by the locative/locomotional semantics of the verb - rather than clausal arguments. In (113), for instance, the NP *kaz* cannot be an argument because the sentence conforms to a monotransitive possessive template in which *yo* and *nĩge* occupy the argument positions, to which a bare locative NP is appended:

(113) *yo nã   te   nĩge   kaz.*
       1s NEG$_{cl}$ have.NPST nobody home
       'I do not have anyone at home.'

Now that we have established that locative phrases (including bare NPs) are not arguments of the verb *te*, we must conclude that there is no structural difference between a presentative clause with a locative adjunct and a locative clause. In other words, DIP makes no formal distinction between locative 'a boy is in my neighbourhood' and presentative 'there is a boy in my neighbourhood'. Sentence (114) is therefore equivalent to both these translations:

(114) *nə   mĩ   bar   te   nə, ũ   rapaz te,   el*
       LOC 1s.POSS neighbourhood EXS.NPST REQ one boy   EXS.NPST 3s
       tə   brĩk-a ku   mĩ.
       IPFV.NPST play-INF COM 1s.OBL

'There is a boy in my neighbourhod/A boy is in my neighbourhood, you know, he plays with me.'

The semantic distinction between the two therefore rests solely on the pragmatic status of the various constituents. The assignment of a presentative or a locative reading in these cases takes into account context, information structure and intonation. It is clear, for instance, that the constituents introduced into the discourse by existentials cannot be definite, as they are new information. A definite referent such as (110a), for instance, therefore favours a locative reading, but indefinite referents make no decisive prediction. An intonational break between the main clause and the locative phrase is an indicator that the clause is to be interpreted as presentative existential and the locative phrase as a secondary, additional element. That seems to be the case in (114), in fact; the locative *nə mĩ bar* 'in my neighbourhood' is separated from the clause *ũ rapaz te* 'there is a boy' by a) intonation, b) preverbal occurrence, and c) a tag[9]. All these elements strengthen the presentative interpretation of this sentence.

7.1.3.4 Possessive clauses

The sentence in example (113) above already hinted at the formal similarities between possessive clauses and existential/locative clauses. The crucial point of contact is the fact that all these constructions make use of the verb *te*. In its possessive function, this verb is transitive, commanding both a possessor NP/pronoun and a possessed NP/pronoun (although ellipsis can intervene here as elsewhere, see 7.8.2). Given their similarities, it is possible to interpret possessive clauses as a type of locative clause construing the possessor as a location. I will return to this issue below.

The possessor NP/pronoun may be bare or it may be case-marked. I will begin by describing the former, as these are the ones which are potentially ambiguous with locative clauses. Consider the following examples:

(115)   a.   *Ryan te   oyt an.*
                R.   have.NPST eight year
                'Ryan is eight years old.' [Lit. 'Ryan has eight years.']

        b.   *yo nã   te   muyt famil.*
                1s NEG$_{cl}$ have.NPST much family
                'I don't have a numerous family.' [Lit. 'I don't have a lot of family.']

        c.   *yo nã   te   niŋũ amig.*
                1s NEG$_{cl}$ have.NPST NEG$_{ct}$ friend
                'I don't have any friends.'

---

[9]The tag corresponds to the particle *nə*, glossed here as a requestative element (REQ). See section 7.6.2.2 for a clarification of its use in polar questions.

In (115), the semantics of the possessed NPs block a locative reading, but (116) is possibly ambiguous[10]:

(116)   a.   *yo te   kaz.*
            1s *te*.NPST home
            'I have a home.'/'I am at home.'

        b.   *yo nã   te   kaz.*
            1s NEG$_{cl}$ *te*.NPST home
            'I don't have a home.'/'I am not at home.'

Context may be enough to disambiguate, otherwise speakers who intend the possessive reading have the option to modify the possessed NP with an indefinite article[11] as in (117a), or an adnominal negator as in (117b):

(117)   a.   *yo te   ũ   kaz.*
            1s *te*.NPST one home
            'I have a home.'

        b.   *yo nã   te   niŋũ kaz.*
            1s NEG$_{cl}$ *te*.NPST NEG$_{ct}$ home
            'I don't have a home.'

Possessor NPs in possessive clauses are often embedded in a PP.[12] It is difficult to ascertain the semantic difference between dative-marked possessor NPs (with one of the Dative prepositions *a* or *pə*) and bare NPs, as this seems to be a matter of free variation.

---

[10] In theory, similar ambiguity may also affect the following sentence:

(1)   *mĩ   saykəl te   kaz.*
      1s.POSS bicycle *te*.NPST home
      'My bicycle is at home.'/?'My bicycle has a home.'

In this case, though, the inanimate nature of the putative possessor *saykəl* renders a possessive interpretation less likely.

[11] The fact that indefiniteness favours the argument interpretation of the NP over a locative reveals that locative reference is expected to be definite. If an indefinite NP constitutes a locative phrase, it will be clearly marked as such, for instance by the prepositions *nə* 'in' or *pə* 'to'.

[12] The occurrence of non-bare possessors in possessive clauses is further evidence that these constructions are modelled on locative clauses and that it is the possessor which constitutes the metaphorical location. It is no coincidence that dative and locative prepositions are selected in these cases; as explained in 7.4.1 below, *direction* falls within the semantic scope of the Dative category. The preposition *jũt də*, meaning 'near' or 'next to' (see 7.4.8), is clearly locative. Many Indian languages show similar locative marking of the possessor in possessive clauses with a similar semantic value (see Masica 1991, Dasgupta 2004, Mahajan 2004, Mistry 2004), which is likely to have provided the model for these constructions in DIP. The present-day use of comitative *ku* in such contexts, rather marginal, is probably an extension due to its functional similarity to *jũt də* as a comitative marker.

Consider the following sentences (118):

(118) a. *may ũ   irmãw yo te.*
          more one brother 1s have.NPST
          'I have another brother.'

       b. *te   bigɔd   pə   el.*
          have.NPST moustache DAT 3s
          'He has a moustache.'

Other possessors are preceded by one of the comitative conjunctions *jũt də* and *ku* instead, the former being widely dominant in such cases. Whenever possessors take the form of comitative PPs, they have a very peculiar meaning; these constructions indicate temporary possession, as exemplified in (119):

(119) a. *mĩ   jũt   nã   te   muyt diɲer nã   te.*
          1s.OBL together NEG$_{cl}$ have.NPST much money NEG$_{cl}$ have.NPST
          'I [Lit. 'with me'] don't have much money.'

       b. *elz vidi   faz-e   kaz   si d-elz jũt   t-iŋ   diɲer.*
          3p IRR.PST make-INF house if of-3p together EXS-PST money
          'They would have made a house if they had the money.'

       c. *ku   el nə   t-iŋ.*
          with 3s NEG$_{cl}$ have-PST
          'He did not have [it].'

See section 9.5.2 for a diachronic account of dative- and comitative-marked possessor arguments.

7.1.4   Conjunct verbs

The predicate units known as *conjunct verbs* (Masica 1991), labelled *nominal-verb complexes* in Cardona and Suthar (2003) are a widely attested and debated characteristic of Indic languages. These complex predicates consist of two lexemes, typically a nominal or adjectival element and a light verb[13]; conjunct verb formation constitutes a very productive verbalising strategy. The defining trait of these complex predicates is that, even

---

[13]A light verb is defined here as a grammatical (rather than lexical) verb 'that functions as a member of a complex predicate construction' (Mohanan 2006:461). Although light verbs are formally equivalent to full verbs, in this function their semantic content is typically altered, such as for example concerning the verb 'to give' in 'to give a gasp'.

though they can easily be segmented into a verbal and a non-verbal parts, they constitute a single syntactic unit; in other words, the non-verbal element does not occupy an argument position. They furthermore differ from copular constructions in that the verbal element is not semantically empty and the conjunct verb can be transitive.

In DIP, the light verb which is the most productive in conjunct verb formation is *faze* 'to make', which combines with both a noun (e.g. *faze fon* 'to make a telephone call' [Lit. 'to make telephone']) or an adjective (e.g *faze lĩp* 'to clean' [Lit. 'to make clean']). The non-verbal elements in these complex predicates are often common DIP words used in other contexts, but they also include borrowed words (e.g. *faze pɛch* 'to squash', with a form of the Gujarati adjective *pechũ* 'squashy'; *faze giligili* 'to tickle', with a truncated and reduplicated form of the Gujarati noun *gəlipəchi* 'tickle'). Occasionally, the additional element may indeed be verbal, as for instance *kima* in *faze kima* 'to roast' - see example (292b) below. The use of conjunct verbs is exemplified for *faze complete* 'to complete' and *faze fon* 'to make a phone call' in (120):

(120) a. *yo t-iŋ   faz-e   complete mĩ   book.*
           1s IPFV-PST make-INF complete 1s.POSS book
           'I was completing my book [i.e. homework].'

       b. *nɔs kwɔn fez   fon   (a elz), elz ain t-iŋ   durmid.*
           1p   when make.PST phone to 3p   3p still IPFV-PST asleep
           'When we called (them), they were still asleep.'

It is true that the verb *faze* may have a causative meaning in combination with verbs - see 7.10.1 for details. In the cases above, however, a causative reading could possibly be assigned to (120a), but not to (120b). *Faze* regularly occurs first in conjunct verbs, but the order may also be inverted as if the remaining element were a regular complement - see section 7.8 for the pragmatic motivation of word order variation. This is exemplified in (121) for the conjunct verb *faze prõt* 'to prepare':

(121) *tɛ   agɔr nə   vey   nad,   tud papɛl jə   prõt jə   fez.*
        until now NEG$_{cl}$ come.PST nothing all   paper already ready already make.PST
        'So far, nothing has come, I have already prepared all the papers.'

Section 9.6.5 explores the link between DIP conjunct verbs and Gujarati conjunct verbs.

7.1.5   Participial constructions

The formation of participles involving the -*d* suffix was described in 6.2, but the issue of their syntactic embedding[14] - and indeed their categorisation as adjectival or verbal forms

---

[14]Passive constructions were never recorded in naturalistic speech, and in elicitation only a few collaborators responded with a construction very close to that of standard Portuguese: the Past form of the

- is somewhat complex. As shown above, *tə/tiŋ* function as verbal auxiliaries, but *te/tiŋ* is also the stage-level copula (see section 7.1.3). The formal dichotomy between auxiliary *tə* and copular *te* is simply based on the fact that the vowel in the auxiliary is often reduced whereas that in the copula is not, but this is tendential rather than deterministic.

The lack of unequivocal formal distinction between auxiliary and copular *tə/te* poses a challenge to determining whether a participle preceded by *tə/te* and *tiŋ* is an adjectival form used predicatively or the head of a VP modified by an auxiliary. The least ambiguous occurrences of participles do suggest that participles enter both constructions, and as such cut across the verb/adjective categorial boundary as defined in 6; in (122), for instance, the participles *foyd* and *brĩkad* are clearly verbal in that they govern a nonfinite complement clause (122a) and a direct object (122b) respectively:

(122) a. õt   te   nə   nɔs t-iŋ   foy-d   ve   theater *i*
           yesterday EXS.NPST REQ 1p IPFV-PST go-PTCP see.INF theater and
           ali   *nɔs vi-w*   Blackmail *sinem.*
           there 1p   see-PST Blackmail movie
           'Yesterday we went to see the theater and there we watched Blackmail.'

       b. *yo time t-iŋ*   *brĩka-d*   computer.
           1s also IPFV-PST play-PTCP computer
           'I also played with my computer.'

Example (123) is interesting because it involves an existential construction:

(123) õt   t-iŋ   ti-d   *mis.*
       yesterday IPFV-PST EXS-PTCP mass
       'There was a mass yesterday.'

On the other hand, the participle in (124) is adjectival, given that the valence of *fika* admits nominals but not verbs:

(124) *nã*   *fik*   *asa-d*   *dret?*
       NEG$_{cl}$ become.NPST roast-PTCP properly

---

auxiliaries *te* or *vay* modifying the participle of the lexical verb and the agent expressed in a *pə*-prepositional phrase:

(1) *ikəl*   *furtĩ foy*   / *t-iŋ*   *mara-d*   *pə purtəgez.*
     DEM$_d$ fort vay.PST   IPFV-PST build-PTCP *pə* portuguese
     'That fort was built by the Portuguese.'

All instances of elicited passives were rather hesitant, which suggests that this is not a normal syntactic possibility in DIP but rather a construction borrowed from standard Portuguese by speakers with a certain proficiency in it. See section 9.6.4 for further discussion.

'Isn't it getting properly roasted?'

All the unequivocally verbal occurrences of participles combine with the Past auxiliary *tiŋ*. It seems that the combination *te* + Participle always constitutes a copular construction. The only potentially ambiguous sentences, therefore, are those involving *tiŋ* and the participle form of an intransitive verb, such as (125):

(125)   a.   *gat t-iŋ   durmi-d.*
             cat IPFV-PST sleep-PTCP
             'The cat was asleep.' or 'The cat slept/had slept.'

        b.   *akəl-ɔr   use nə   sab,   use nə   t-iŋ   nasi-d.*
             DEM$_d$-hour 2   NEG$_{cl}$ know.NPST 2   NEG$_{cl}$ IPFV-PST be.born-PTCP
             'You don't know about that time, you weren't born / you hadn't been born.'

It is clear, at any rate, that the participial form of a verb is charged with perfective semantics. Despite the fact that *tiŋ* is prototypically a marker of imperfectivity, examples (122) and (123) testify to the fact the the *tiŋ* + Participle sequence has a Past perfective - rather than pluperfect - reading. The difference between a periphrastic Past perfective of this type and an inflectional Past perfective is subtle and not always easy to pinpoint. In this respect, DIP appears to mirror several Indic languages in that when the periphrastic Past perfective is employed, a subsequent action is expected. This subsequent action may be expressed, as in (122a), or not, as in (122b).

The occurrence of verbal forms which look like gerunds, recognisable by the suffix -*n* is interesting and somewhat complex. In comparable contexts - particularly to express progressive actions -, some speakers were seen to use gerunds recurrently where others consistently employ an infinitival verb form. As it turns out, most instances of these forms in the corpus occur in the speech of people who are either natives of Daman or direct descendants of (a) Damanese. The regular use of gerunds in modern-day Daman Indo-Portuguese is well documented (Clements and Koontz-Garboden 2002), and as such this pattern of use in Diu is not surprising. A more complete description of this phenomenon, and the implications of the observed family effects, can be found in section 2.2.2.

Despite these considerations, there is a small number of crystallised gerunds used across the board in DIP, some of which are given below in (126). The expression *fika saben* 'to find out, to hear, to be told' differs from the simple use of the verb *sabe* 'to know' in that the former underscores a process by which knowledge is attained.

(126)   a.   *yo jə   fik-o   sabe-n.*
             1s already become-PST know-PROG
             'I've already heard/found out.'

b. *Austin, yo kɛ   vay   anda-n.   anda-n   yo kɛ*
   A.   1s want.NPST go.INF walk-PROG walk-PROG 1s want.NPST
   vay.
   go.INF
   'Austin, I want to walk. I want to *walk*.'

For most speakers of DIP gerundive -*n* is not a productive suffix unless there is a strong tradition of speaking Daman Indo-Portuguese in the immediate family.

7.2 The noun phrase

The basic structure of the NP, introduced in section 6.5, is repeated here for convenience:

(127)   Dct + Qtf/Ord + [Int + Adj] + N + PP/Rel

This scheme is an abstraction based on the observation of several shorter NPs, given that no NP in the corpus ever fills all of these slots. Some relevant examples are given in (128) as illustration - in this section, the relevant NPs are isolated in square brackets in example sentences, whenever necessary:

(128)   a.   [*es   tud papɛl*] *de-w.*
          DEM$_p$ all paper give-PST
          '(I) gave all the papers.'

     a.   [*es   ũ   rapaz*] *te   aki.*
          DEM$_p$ one boy   EXS.NPST here
          'This one boy is here.'

     b.   [mĩ tud amig]
          1s.POSS PL friend
          'my friends'

     b.   *Leslie ɛ   *[*bẽy piken baba*] *ɛ.*
          L.   COP$_i$.NPST very small baby COP$_i$.NPST
          'Leslie is a very small baby.'

The NP-negator (NEG$_{ct}$, see section 7.7) and adnominal interrogatives (AdnInt), on the other hand, do not co-ocur with other nominal modifiers. The structure of NPs containing these words is that given in (129):

(129) NEG$_{ct}$/AdnInt + N

Minimally, an NP consists simply of its nominal head but, as indicated in (127) and (129), it admits a host of modifiers. The functions of NP operators is diverse; for this description, I will follow Rijkhoff's (2002) typology which classifies NP modifiers as Qualifying (if they specify more or less inherent characteristics of the referent), Quantifying (those which attribute quantitative measurement to the referent) and Localising Modifiers (the ones that locate the referent, not simply in physical terms but within the discourse world, such as discourse deictic markers). I will discuss these various domains in turn, before describing other NP operations such as number, definiteness marking and negation.

7.2.1 Qualifying modification

NP-internal qualification of the DIP noun is typically achieved by means of an Adjective Phrase (for which, see section 7.3 below) headed by an adjective (130a). The language also resorts to what may be interpreted as appositional modification (130b), in which a noun is made to modify another[15] - see also the discussion in 6.4:

(130) a. *ŭ piken rat   t-iŋ   vay nə d-el kamiŋ.*
          one small mouse IPFV-PST go.INF LOC of-3s way
          'A small mouse was going about his own way.'

       b. *ali   nɔs vi-w   Blackmail sinem.*
          there 1p see-PST Blackmail film
          'There, we saw the film *Blackmail*.'

The canonical position of the attributive AdjP is prenominal, as illustrated above. Examples of postposed AdjPs are few and far between, exemplified by the position of *vɛy* in (131a) and *pɔb* in (131b):

(131) a. *de-w   rəkri   de-w,   dəpəy de-w   kader, fugãw vɛy.*
          give-PST food.stall give-PST then gave-PST chair stove old
          '(He) gave (me) the food stall, and also chairs, an old stove.'

       b. *nã   lɛv   mal, jẽt   pɔb.*
          NEG$_{cl}$ take.NPST badly people poor
          'Don't feel offended [Lit. 'Don't take [it] badly'], [we are] poor people.'

---

[15] In many languages, non-restrictive relative clauses simply add relevant information concerning the referent, and as such they can be seen as qualifying modifiers. In DIP, however, all naturally-occurring relative clauses in the corpus are restrictive, and as such qualify as localising modifiers instead (see 7.2.3). The only non-restrictive relative clauses recorded were the result of elicitation.

The preposition *də* functions as a generic relator, allowing elements other than AdjPs to modify a noun. This is fully expanded in section 7.4.3.

7.2.2 Quantifying modification

Quantification is marked primarily by the elements classified as quantifiers in 6.5.2: these include lexical quantifiers and cardinal numerals. They precede both the head noun and the AdjP, although it is rather uncommon to find numerals and adjectives in sequence (mostly due to the fact that adjectives are more common in predicative position than attributive position, see 7.3).

(132)   a.   *dəpəy də sĩk/ bastãt an.*
             after of five many year
             'After five/many years.'

       b.   *dəpəy põy   pok sum də limãw.*
             then   put.NPST little juice of lime
             'Then you add a little lime juice.'

       c.   *yo t-iŋ   doy võtad.*
             1s have-PST two wish
             'I was of two minds.' [Lit. 'I had two wishes.']

Quantifiers with absolute quantifying semantics, such as the universal quantifier *tud* 'all' in its adnominal use and the NP-negator *niŋũ* 'no [X]' - but not the negative indefinite pro-forms *nĩge* 'nobody' and *nad* 'nothing' - can also be interpreted as special types of quantifying modifiers:

(133)   a.   yo tud koyz sab faz-e.
             1s all thing know.NPST make-INF
             'I can do all things.'

       b.   yo niŋũ koyz nə sab faz-e.
             1s NEG$_{ct}$ thing NEG$_{cl}$ know.NPST make-INF
             'I can do no things.'

The expression of number within the NP is another function of quantifying operators. Plurality is marked periphrastically in DIP with resort to the word *tud*, which adds to its function as the universal quantifier (both adnominal and pronominal) that of a plural marker. Given that number marking is a somewhat complex issue, it will be assigned a separate descriptive section - see 7.2.7 below.

It is hardly surprising that a given word might simultaneously function as a quantifier and a plural marker, since both functions result in proper quantification of the NP. It is more unexpected, however, that a quantifying element should accumulate a localising function; such is the case of the numeral *ŭ* 'one', which also functions as an indefinite marker. See section 7.2.6 for a full account of the grammatical use of *ŭ*.

7.2.3 Localising modification

In order to grasp the concept of localising modifiers, consider the function of ordinal numerals, which are extremely rare in DIP, as explained in 6.5.3 and 8.4.2. Their function is substantially different from that of cardinal numerals because, rather than quantifying, ordinals locate the referent within a shared or assumed chain of possible referents.

(134) *tɛrser kaz   dəpəy də igrej.*
       third   house after of church
       'The third house counting from the church.'

Likewise, possessives and determiners all contribute to anchor the referent within discourse and/or deictic reality. In DIP, attributive possession may be conveyed in two ways: through a possessive preposed to the head noun plus any qualifying and quantifying modifiers (135a), or through a postnominal *də*-PP (135b):

(135) a. *d-use gran amig   kẽ   ɛ?*
           of-2   big   friend who COP$_i$.NPST
           'Who is your best friend?'

       b. *el foy   kaz   də irmã.*
           3s go.PST house of sister
           'He went to the house of his sister.'

Notice how, with the exception of the forms for first person singular and plural, the forms of the possessives are formally equivalent to a *də*-PP containing a personal pronoun (see section 6.5.4). Although it is not abusive to interpret these as lexicalised possessives, given the observed phonological bleaching of the putative preposition, the structure of these compounds is not lost on the speakers.

Despite these general tendencies, there is room for some variation. Very rarely, the *də*-PP type of structure of 2nd and 3rd person possessives is extended to first person, as in (136):

(136) *yo ku   də mĩ   amig, nɔs t-iŋ   i-d   ali   nə Vanakbara.*
       1s with of 1s.OBL friend 1p IPFV-PST go.PTCP there LOC V.

'Me and my friend(s), we went to Vanakbara today.'

Occasionally, this structure occurs in postnominal position, if embedded in *də*-PP[16], as in (137):

(137) ũ   irmãw də mĩ   fik   Go.
       one brother of 1s.OBL stay.NPST Goa
       'A brother of mine lives in Goa.'

Proximal and distal demonstratives occur in absolute first position in the NP. Their prototypical function is that of localising a referent with regard to a given deictic point, which is often the location of the speaker. In DIP, demonstratives code a twofold distance contrast; the proximal (DEM_p) *es* 'this' indicates that the referent is in the vicinity of the speaker and distal (DEM_d) *ikəl* 'that' locates it away from the speaker. The particle *ot(r)* 'other', which indicates the absence of the referent from the deictic/discourse context, may either be phrase-initial or follow another locative modifier. The use of these modifiers is demonstrated in the following sentences:

(138) a. es raprig a fika ũ bunit muyɛr.
           DEM_p girl IRR.NPST become.INF one beautiful woman
           'This girl will become a beautiful woman.'

       b. *dig   a   mĩ   kẽ   de-w   a   use ikəl   anɛl*
           tell.NPST DAT 1s.OBL who give-PST DAT 2 DEM_d ring
           'Tell me who gave you that ring.'

       c. *mĩ   irmãw te   nə   ot~ot   iskɔl.*
           1s.POSS brother EXS.NPST LOC other~other school
           'My brother is at a different school.'

---

[16]Notice that, in predicative position, pronominal possessors are only allowed if embedded in a *də*-PP:

(1) *nãw, es   ɛ   də mĩ   mem   ɛ.*
     NEG_p DEM_p COP_i.NPST of 1s.OBL EMPH COP_i.NPST
     'No, this is my own.'

An additional complication of the syntax of possession arises from the fact that a small number of short *də*-PPs have been recorded in prenominal rather than postnominal position (2), although it must be stressed that this is by no means the rule - see 7.4.3 for further details:

(2) *də tɛtɛ kaz   jə   bẽze-w?*
     of T.   house already bless-PST
     'Has [he] already blessed Tɛtɛ's house?'

169

On other occasions, the deictic modifiers do not locate the referent within the physical world but rather in discourse (i.e. discourse deixis), referring anaphorically to previously mentioned elements. Both demonstratives can be used in this function, in which the proximal/distal distinction appears to be irrelevant:

(139) a. *i dəpəy es   leopard foy   murd-e pə   lion.*
           and then DEM$_p$ leopard go-PST bite   DAT lion
           'And then the leopard went to bite the lion.'

       b. *vi-w   n-ikəl   di kẽ t-iŋ   mĩ   kaz, sɔ ikəl   tăt*
           see-PST in-DEM$_d$ day who EXS-PST 1s.POSS house only DEM$_d$ many
           piso.
           person
           '(You) saw who was in my house on that day, only that many people.'

The index of reference for *ot(r)* may also be entirely discourse-internal, as shown in (140):

(140) *oj   pɔd   diskãs-a, ot   di də sɛman nə   kusɛg,   nã*
       today can.NPST rest-INF other day of week   NEG$_{cl}$ can.NPST NEG$_{cl}$
       apaŋ   tẽp.
       find.NPST time
       'Today (I) can rest, other days of the week (I) don't manage, (I) don't find the time.'

In some cases still, both deixis and anaphoric reference fail to account for the selection of demonstratives, and this reveals an additional function of these forms as markers of definiteness. (In)definiteness marking is indeed a prototypical function of localising modifiers within the NP but, in order to fully capture its complexity, the issue is separately treated in section 7.2.6.

Restrictive relative clauses embedded in the NP are also important localising modifiers, as they enhance the referential accuracy of the NP. The prototypical position of the relative clause is after the relativised element, as in (141):

(141) *tud koyz ki   use dew   te   nə mĩ   kaz.*
       all thing REL 2   give.PST EXS.NPST LOC 1s.POSS house
       'All things that you gave me are in my house.'

This order can also be reversed in response to the pragmatic status of the various constituents, as shown in (142).

(142) *pɛrt də kapɛl ki te ikəl kaz?*
      near of chapel REL EXS.NPST DEM$_d$ house
      'The house which is next to the chapel?'

    In the previous example, one will notice that the relative element does not occur in the beginning of the relative clause; the internal structure of relative clauses admits various configurations, for further information see section 7.10.3.

7.2.4 Adnominal interrogatives

In contrast to the bulk of the DIP question words, which are pro-forms (see 6.6.3 and 7.6.2), three interrogative elements routinely appear in noun modifier position (see also 6.5.5). These typically demand either quantification or referential clarity. Quantification is elicited with the question word *kwõt* 'how much/many', and *kwɔl* 'what, which' (occasionally also *ki*) is a more generic interrogative. The distinction between an adnominal and a pronominal interrogative is explored in (143). As explained in 6.6.3, the form *uki* 'what' is exclusively pronominal; *ki* 'what', on the other hand, occurs both within an NP and as a pronominal interrogative equivalent to *uki*. Additionally, *kwɔl* 'what, which' is only allowed in adnominal position. In example (143a), *ki* and *kwɔl* appear as nominal modifiers, while *uki* is blocked. In (143b), the interrogative stands for an entire NP, so both *ki* and *uki* are allowed; *kwɔl* is classified as infelicitous rather than outright ungrammatical simply because it may be allowed to stand for an entire NP from which the head noun has been elliptically removed:

(143) a. *use ki/*uki/kwɔl trabay a faz-e?*
            2   what   work IRR.NPST do-INF
            'What job will you do?'

      b. *use ki/uki/?kwɔl a faz-e?*
            2   what   IRR.NPST do-INF
            'What will you do?'

    The head of the NP is usually expressed, as in (144a), but may be left out whenever it is clear from context, as in (144b) - for a discussion of this matter in connection with observed patterns of ellipsis in DIP, see section 7.8.2:

(144) a. *use kwõt an te?*
            2   how.much year have.NPST
            'How old are you?' [lit. 'How many years do you have?']

      b. *ũ istɔr. . . kwɔl a kõt-a?*
            one story   which IRR.NPST tell-INF

'A story... Which (one) should I tell?'

In consonance with the schematic representation given under (129), the corpus does not contain any instances of the co-occurrence of adnominal interrogative forms and other noun modifiers. The only exceptions involve a postnominal relator *də*-PP, i.e. one allowing the modification of a nominal head by another nominal head (see 7.4.3). This is exemplified in (145):

(145)   *kwõt   garaf də* beer *t-iŋ   nə   chãw?*
        how.much bottle of beer EXS-PST LOC ground
        'How many bottles of beer were there on the floor?'

Further examples of adnominal interrogative constructions can be found in section 7.6.2.3, which describes the syntax of content questions in DIP. See also 9.6.1.

7.2.5   Noun phrase negation

The NP-negator *niŋŭ* 'no [X]' is distinct from the universal negators *nĩge* 'nobody', *nad* 'nothing' and *nuk* 'never' in that it is never pronominal; the universal negators do not occur in the presence of a nominal head, although they can be replaced with semantically equivalent negated NPs: e.g. *niŋŭ piso* 'no person' for *nĩge*, *niŋŭ koyz* 'no thing' for *nad*, and *niŋŭ vez* 'no occasion' for *nuk*.

The NP-negator is preposed to the head noun, as shown in (146). This element is glossed NEG_{ct}, i.e. constituent negator, to distinguish it from the clausal negator (NEG_{cl}) as well as the sentential negator (NEG_p) - see section 7.7 for further details.

(146)   a.   *tɛ   agɔr niŋŭ   rɛpos nã   de-w.*
             until now NEG_{ct} answer NEG_{cl} give-PST
             'So far, (they) have not given any answer.'

        b.   *nã   da   võtad də ir   niŋŭ   part.*
             NEG_{cl} give.NPST wish of go.INF NEG_{ct} place
             'One has no wish to go anywhere.'

Negated NPs may perform various functions. In (146a), *niŋŭ rɛpos* is a direct object, whereas in (146) the NP *niŋŭ part* functions as an indefinite adverbial. These observations also make it clear that the negative polarity value of the NP-negator does not invalidate clausal negation; for further information on the phenomenon of negative concord in DIP, see section 7.7.

Similarly to adnominal interrogatives, the NP-negator *niŋŭ* does not co-occur with other noun modifiers (but see example (45) and discussion thereof).

172

7.2.6 Definiteness

As mentioned in section 7.2.3, both demonstratives *es* 'this' and *ikəl* 'that' function as definite articles, which is particularly evident if the referent they modify cannot be retrieved from the deictic space or discourse. In (147), the referent *raprig* 'girl' is absent from the physical environment of the utterance, and it has not been introduced into the discourse before. The demonstrative constructs definiteness, in this case, based on (assumed) shared knowledge between the speaker and the interlocutor.

(147) *es   tud ɔn   foy   raprig?*
       DEM$_p$ PL where go.PST girl
       'Where did the girls go?'

In (148) below, the referent *papɛl* had previously been introduced into the discourse, so that the reference is unequivocal:

(148) *es   tud papɛl de-w   may də dez an jə   fik-o.*
       DEM$_p$ PL paper give-PST more of ten year already become-PST
       '(I) submitted these papers more than ten years ago.'

Such cases occupy the border between abstract definiteness marking and anaphora, but the fact remains that in either case the referent is construed as definite.

The numeral *ũ* 'one', on the other hand, functions as an indefinite article. In truth, any quantifier produces indefinite reference unless the NP is otherwise marked for definiteness, but *ũ* is used by default whenever the referent is not known to the interlocutor (when it is first introduced, for instance) and quantification is not particularly relevant:

(149) *ali   ve   ũ   dɛmɔn si sa-i.*
       there see.NPST one ghost if leave-INF
       'Watch out if a ghost will come out.'

The DIP bare noun gives no indication concerning definiteness, and both the definiteness and indefiniteness markers are optional. In that sense, these grammatical agents are definiteness *disambiguators*. As an illustration of the optionality of (in)definiteness markers, consider example (150):

(150) *i   dəpəy gat foy   fal-a   pə kãw, i kãw ku   lion kuməs-o.*
       and then cat go.PST speak-INF DAT dog and dog COM lion begin-PST
       jug-a.
       play-INF
       'and then the cat went and told the dog, and the dog started playing with the lion.'

All the referents (*gat*, *kãw* and *lion*) are definite, because they have already been introduced in the story, but they are still represented by bare nouns.

7.2.7 Number

Number marking is a function of quantifying operators of the NP (see section 7.2.2). DIP recognises two number categories, viz. singular and plural, and it is plurality - not singularity - which may be overtly marked. This is not to say that unmarked nouns are singular; in fact, according to the typology of nouns proposed in Rijkhoff (2002), DIP nouns must be classified as *set nouns*, which means they are entirely unspecified concerning number - this classification is discussed in more detail in 6.4. As such, the marker of plurality is more accurately a *plural disambiguator*, or *collectiviser*.

Two different types of plurals can be recognised in DIP, viz. the *additive plural* and the *similative plural*, which differ in that the former indicates a homogeneous set made up of identical elements while the second indicates a group of heterogeneous referents. Both kinds of plural are marked by the particle *tud*, identical to the universal quantifier, but differ in their syntactic configuration. These will be described in turn.

7.2.7.1 Additive plural

Additive plurals indicate sets of referents in which all the participants are identical, i.e. all can be referred to by means of the pluralised NP. The English phrase *my brown shirts* is an example of an additive plural in that all members of the set are characterised as being shirts, brown, and belonging to the speaker.

In DIP, additive plural marking does not co-occur with other means of quantification such as numerals. It is always optional but strongly favoured by definiteness[17], the reason for which pluralising *tud* overwhelmingly occurs associated to a demonstrative[18] (which expresses definiteness, see 7.2.6) as in (151a). The plural marker also co-occurs with other NP modifiers(151b), and there are recorded instances of pluralised demonstrative pronouns (151c):

(151) a. *ikəl   tud koyz ki   lεv-o   nə   museum father Marian ki*
           DEM$_d$ PL thing REL take-PST LOC museum father M.   REL

---

[17]Generic referents are usually not marked for plurality. Consider the following example:

(1) *kriãs dig   "es kumid ε bunit " may nɔs dig   "es kumid ε muyt gustoz "*.
     child say.NPST –   but 1p say.NPST –
     'Children say "*es kumid ε bunit*" but we say "*es kumid ε muyt gustoz*"'.

[18]The fact that *DEM + tud* apparently constitute a full participant in (151c), as well as (152) below, may be interpreted in two different ways: either a) the absence of a nominal head is the result of ellipsis, or b) in its pluralising function *tud* actually preferentially attaches to an indicator of definiteness, with which it constitutes a syntactic unit.

> *lɛv-o,   nə?*
> take-PST REQ
> 'Wasn't it father Mariano who took those things to the museum?'

    b. *mĩ̃   tud amig*
> 1s.POSS PL friend
> 'my friends'

    c. *es   tud nã   te   nə   Diw, es   ɛ   də Gogla.*
> DEM$_p$ PL NEG$_{cl}$ EXS.NPST LOC Diu DEM$_p$ COP$_i$.NPST of Goghla
> 'These do not live in Diu, they are from Goghla.'

The semantic similarity between a universal quantifier and an additive plural marker is evident, and this is indeed a well attested path of grammaticalisation crosslinguistically - see Heine and Kuteva (2002). Given that, in syntactic terms, the word does not behave significantly differently in both functions, it is often difficult to tell them apart. Two factors justify the proposed macrofunctional interpretation of *tud*. The first of these is that the occurrence of this element is much more frequent than what might be expected if universal reference were its sole dimension. In addition, elicited data makes it clear that, when asked to pluralise an NP, speakers will resort to *tud*, as seen in the pair of elicited sentences below:

(152) a. *es   tud ɛ   kaz   də pad.*
> DEM$_p$ PL COP$_i$.NPST house of priest
> 'These are the houses of the priest.'

    b. *es   tud ɛ   kaz   də tud pad.*
> DEM$_p$ PL COP$_i$.NPST house of all priest
> 'These are the houses of the priests.'

The corpus contains rare examples of nominal reduplication with a pluralising function. This is a very marginal strategy in modern DIP. Consider the following example:

(153) *dəpəy muyɛ~muyɛr   də Manu, muyɛ~muyɛr   də ɔrlãd.*
> then woman~woman of M.   woman~woman of O.
> 'Then the women of Manu('s family), the women of Orlando('s family).'

The pluralising function of nominal reduplication is clear in (153)[19], but less so in (154) below because nominal reduplication also often imparts a distributive meaning (see section 6.4):

---

[19]The occurrence of *mulher mulher* in the Indo-Portuguese variety of Diu was already reported in Dalgado (1903).

(154) es   tud rəkri   tud ɛ   də moyr∼moyr.
       DEM$_p$ all food.stall all COP$_i$.NPST of muslim∼muslim
       'All these food stalls belong to Muslims.'

The sentence in (153) was produced by an elderly lady, whereas (154) was uttered by a middle-aged man whose parents are credited by others as having spoken a particularly 'old' variety of DIP. This distribution ties in well with the fact that pluralising nominal reduplication used to be more prevalent in past stages of the language; the late 19th-century corpus of DIP recorded and studied in Schuchardt (1883), as well as the data adduced by Quadros (1907:192*ff*), indicate that reduplication was then the standard pluralising strategy.

7.2.7.2   Similative plural

The similative plural is a special type of plural reference distinct from the additive plural by two semantic traits, viz. reference to a heterogeneous group and reference to a group defined as having features similar to a given prominent member. The similative plural also differs slightly from the so-called *associative plural* because of the relationship both establish between the members of the set of referents. In associative plurals, such as for instance the Japanese phrase *Tanaka-tachi* 'Tanaka and his associates' (Daniel and Moravcsik 2005:150) or Portuguese *os meus pais* 'my parents' [Lit. 'my fathers'], the various referents share some type of formal association (Moravcsik 2003). In similative plurals, on the other hand, reference is restricted to entities which share some kind of undefined similarity with a given standard, which in a sense contains but also transcends the stricter semantics of the associative plural.

In DIP, *tud* marks both the similative and the additive plural, but its position in relation to the head noun is different. Whereas *tud* occurs prenominally in additive plurals, as described in 7.2.7.1, it is postnominal in similative plural constructions. Some examples of the latter follow:

(155)   a.   *el t-iŋ   vay   nə   ŭ* jungle *pu   traz-e   koys, aɾos tud.*
             3s IPFV-PST go.INF LOC one jungle PURP bring-INF thing rice SIML
             'He went into a jungle to bring some things, rice and all that.'

        b.   *nã   apaŋ   tĕp, muyt trabay tud, dəpəy fik   kãsad.*
             NEG$_{cl}$ find.NPST time much work   SIML then   become.NPST tired
             '[I] have no time, a lot of work and all that, then I get tired.'

        c.   *asĩ kwɔn vĕy   mĩ nitiŋ   tud,   nɔs fal   ikəl   mem*
             so   when come.NPST my grandchild SIML 1p speak.NPST DEM$_d$ EMPH
             ẽ kaz   purtəgez   kom nɔs fal   tud di.
             in house portuguese like 1p speak.NPST all day

176

'So, when my grandchildren or so come, we speak Portuguese at home like we do every day.'

7.2.8 Discontinuous NPs

Apart from being an instance of similative plural marking, example (155c) above has another interesting characteristic, namely that the NP *ikəl mem purtəgez* appears interrupted by the locative phrase *ẽ kaz*. The occurrence of discontinuous NPs is not infrequent in DIP. In such instances, canonical NPs are interrupted by an element external to it, and this may or may not involve an alteration of the normal order of noun modifiers. An example is given in (156):

(156)   *es   tud ɔn   foy   raprig?*
        DEM$_p$ PL   where go.PST girl
        'Where did the girls go?'

In (156), the NP *es tud raprig* encloses the predicate and interrogative pro-form *ɔn foy*, and yet - disregarding the discontinuity - the order of the NP constituents is canonical: DEM + Qtf/PL + N. In (157) below, on the other hand, the adjective *pikənin* occurs not only segmented from the NP but also postnominally:

(157)   *ali   ikəl   gate nã   te   apərs-e   pikənin?*
        there DEM$_d$ gate NEG$_{cl}$ IPFV.NPST show.INF small
        'Doesn't that small gate show over there?' [i.e., 'Do you see that small gate over there?']

It is not clear what determines the discontinuous occurrence of NPs, although one must entertain the possibility that the latter segment of the NP constitutes an afterthought. In either of the instances reported here, intonation is not decisive, as there is no unequivocal prosodic break to suggest a clausal boundary.

A different but related phenomenon is the discontinuous occurrence of conjoined NPs. Two examples are given in (158):

(158)   a. *ŭ   makak t-iŋ   i   ŭ   crocodile.*
            one monkey EXS-PST   and one crocodile
            'There was [once] a monkey and a crocodile.'

        b. *trey irmãw te   nə   mĩ   Go i   ŭ   irmã.*
            three brother have.NPST REQ 1s.OBL Goa and one sister
            'There are three brothers of mine and one sister in Goa, you know.'

Example (158a) is rather straightforward, one can see that the conjoined NPs are separated by the main verb. The same is true of (158b), but there are certain odd syntactic elements in this sentence, such as the hesitant addition of the oblique pronoun *mĩ*. This may be interpreted either as a displaced possessive modifier of the NP, or as a truncated Dative-marked possessor (see section 7.1.3.4). The fact remains, in any case, that the conjoined NPs *trey irmãw* and *ũ irmã* are mediated by syntactic material external to any of the NPs.

7.3   The adjective/adverb phrases

Adjective Phrases (AdjP) and Adverb Phrases (AdvP) have identical internal structures, although they differ substantially in distributional and functional terms - for a motivation of the adjective/adverb divide in DIP, see section 6.3. These phrases are headed by an adjective or an adverb and may contain an optional preposed intensifier (see chapter 6 fn. 18 in section 6.5.2 for a definition of the subclass) which gradates the semantics of the head and which is itself subject to modification.

Adjective phrases containing a modifier preferably appear in predicative position, but also within NPs; both syntactic positions are demonstrated in (159)[20], two sentences produced by the same speaker within the same conversation:

(159) a. *asĩ   ε   muyt gran nom.*
           thus COP$_i$.NPST very big   name
           'It is a very big name.'

       b. *mĩ   nom   ε   muyt kõprid, nə?*
           1s.POSS name COP$_i$.NPST very long   REQ
           'My name is very long, isn't it?'

Apart from copular constructions, AdjPs may occur independently from an NP as complements to certain verbs such as *fika* 'to become, to remain' or *aprəse* 'to seem' (cf. section 7.11.2 below). These verbs are not treated as (semi-)copulas in this study because they impose certain selection restrictions (e.g. *aprəse* does not accept a nominal complement without some additional grammatical marking). The following example demonstrates the occurrence of an adjective as the complement of the verb *fika* - see also (138) above:

(160) *el fik-o   ĩtrujad i   el foy   kas.*
        3s become-PST fooled and 3s go.PST house
        'He was fooled and he went home.'

---

[20] In (159a), there is no significant intonational break to suggest *nom* to be an afterthought addition. It is still admissible, however, that *nom* and *muyt gran* constitute separate components of a copular predication with non-prototypical word order.

The head of an AdjP may be compound, as happens with the English expression *blue colour* in (161):

(161) *ikəl*   blue colour *kaz*   *nã*   *tə*   *apərs-e?*
       DEM$_d$ blue colour house NEG$_{cl}$ IPFV.NPST be.visible-INF
       'Do you see that blue(-coloured) house? [Lit. Isn't that blue(-coloured) house visible?]'

An example of an adverb phrase containing an intensifier can be found in (165b), in section 6.7. Another one, which contains a modified intensifier, is given in (162):

(162) *ãta*   *nɔs fal*   *pok*   *mays* melhor *asĩ,*   *ẽ*   *kaz*   *nɔs fal*
       then 1p speak.NPST little more *melhor* thus in house 1p   speak.NPST
       rough~rough.
       rough~rough
       'Then we speak a little bit better, at home we speak very rough.'

This sentence deserves a number of comments. On the one hand, this is one of the few corpus occurrences of the Portuguese superlative form *melhor* 'better'. It also features the adverbial use, without any formal change, of the English adjective *rough* which, in addition, is reduplicated for intensification; this reveals yet another similarity between adjectives and adverbs, namely that reduplication has an intensifying effect (see sections 6.3 and 6.5.1). But the crucial point it makes, for this section, is that *mays* 'more' is modified by the intensifier *pok* 'a little'. Inasmuch as the quantifiers *may(s)* 'more', *men* 'less' and *tãt* 'as [much] as' (which intrinsically contrast the element they modify to a given standard) enter the structure of AdjP and AdvP, they can be interpreted as *comparative intensifiers* (see 7.11). Examples like those in (163) clarify their function as intensifiers:

(163) a. *i*   *rey* *də* jungle *fal-o*   *use* *ɛ*   *tãt piken use uki*   *a*
           and king of jungle say-PST 2   COP$_i$.NPST so small 2s   what IRR.NPST
           faz-e?
           do-INF
           'And the king of the jungle said: "you are so small, what will you do?".'

       b. *may tras vay, may may may.*
           more back go   more more more
           'Go further back [lit. go more back], further, further, further.'

It is interesting to notice that a prepositional phrase functionally and semantically equivalent to an adverb may be treated as an adverb and occupy the head of an AdvP. Consider the following sentence, in which the instrumental phrase *ku kuydad* functions as a manner adverbial:

179

(164)   *nɔs kõvɛrs   ǔ   pok asĩ̌ mays ku kuydad.*
        1p talk.NPST one little thus more INS care
        'We talk a bit more carefully [Lit. 'a little more with care'].'

Similarly, certain adverbial expressions consist of a noun preceded by the preposition *də*, such as for instance *də maŋã* 'in the morning' and *də fors* 'fast'[21]:

(165)   a.   *oj   də maŋã   yo ko   Fabian t-iŋ   gi-a   saykəl.*
             today of morning 1s COM F.   IPFV-PST ride-INF bicycle
             'Today in the morning me and Fabian rode the bicycle.'

        b.   *bɛ̌y də fors   foy.*
             very of strength go.PST
             '[He] went very fast.'

In the corpus, the only element reported to occur in between the intensifier and the head of an AdjP or AdvP is the focus particle *mem* - see 7.8.5 for clarification.

7.4   Prepositional phrases

Prepositional Phrases (PPs) are governed by an adposition (see 6.7), which in DIP precedes the constituents (typically NPs and pronouns) within its scope - with a few exceptions allowed, for which see sections 7.4.8 and 7.4.10.

One important characteristic of some temporal and causal PPs is that they may contain entire finite clauses, provided that the preposition and the clause are mediated by a relative pro-form *ki*, with an implied, unexpressed nominal. nonfinite verb forms, however, being close in many respects to nominal forms, do not require this construction. Consider, as illustration, the examples provided in 7.4.2 for *atɛ* 'until' and in 7.4.10 concerning the complex preposition *ãt də* 'before'.

PPs constitute not only adjuncts of the clause but also arguments, given that prepositions are largely responsible for case marking in DIP. This section describes in detail the functions and characteristics of PPs, organised according to the various adpositions. In the interest of clarity, PPs introduced by equivalent or related prepositions are clustered under one single heading.

7.4.1   *A/pə*

A and *pə* typically indicate Dative case, which is interpreted here as a polysemous category. The exact semantic value of datives has been the matter of some debate. One approach has

---

[21]In view of these examples, one might attribute an adverbialising function to the prepositions *də* and *ku*, although this is not productive beyond a few cases.

been to attempt a broad, minimal characterisation - a *monosemist* approach as opposed to a *polysemist* one, in Haspelmath's (2003) words - such as, for example, Van Hoecke's (1996:31) definition of the Latin dative as a case 'which serves as the limit of the predicate in the sense that it indicates the ultimate term towards which the action or process referred to ends'. One might also prefer to atomise the subcategories associated with the Dative, therefore recognising a variety of possible functions, not all of which are selected by dative markers in all languages. It will become clear from the present section, as well as 7.5 below, that the DIP Dative markers apply not simply to recipient but also experiencer and beneficiary arguments. The prepositions further mark direction and purpose adjuncts.

These two prepositions are be equivalent but differ in their distribution: in general, the form *a* marks pronominal arguments, while *pə* (occasionally realised *pu*) is used with every other NP. This is a very strong distributional rule, although counterexamples do occur. The regular distribution of these two forms is exemplified in (166), marking the recipient argument of the ditransitive verb *dize* 'to tell':

(166)   a.   *dəpəy el foy   i   el dis   pə   irmã* [...]
             then 3s go.PST and 3s say.PST DAT sister
             'Then he went and he told his sister [...]'

        b.   *yo nã   kɛr   kõnt-a, dig   a   el.*
             1s NEG$_{cl}$ want.NPST tell-INF say.NPST DAT 3s
             'I do not want to tell, ask him.'

This pattern of case marking is not exclusively reserved for beneficiaries and recipients. In DIP, these prepositions also mark patients (167a) as well as the most agent-like arguments of experiencer verbs (167b,c) - i.e., those which are outside the control of the arguments[22]:

(167)   a.   *ikəl   lion vey   i   rasp-o   pə   gat.*
             DEM$_d$ lion come.PST and scratch-PST DAT cat
             'The lion came and scratched the cat.'

        b.   *a   mĩ   nã   apaŋ   tẽp.*
             DAT 1s.OBL NEG$_{cl}$ catch.NPST time
             'I don't find the time.'

        c.   *a   el te   med.*
             DAT 3s have.NPST fear
             'He is scared.'

---

[22] These patterns of dative assignment are firmly established in DIP. Additionally, there are other domains in which there is some variation between the various speakers. For more information, see section 7.5.

The sentences in (167b,c) are examples of what has come to be known as *Dative Subjects* in the literature on South Asian languages. Apart from subjects of sensory verbs, dative subjects also occur in DIP when the verb expresses an external circumstance outside the control of the subject:

(168)　*a*　*mĩ*　*təme sẽt*　*así*　*kom vɛrgoŋ purke*　*nɔs ẽ kaz*　*fal*　*así*
　　　　DAT 1s.OBL also feel.NPST thus like shame because 1p in house speak thus
　　　　nə?
　　　　REQ
　　　　'I also feel a bit of shame because at home we speak like this, you see?'

An entirely different use of *pə* is to indicate direction, with motion verbs. This is exemplified in (169):

(169)　*sab*　*nɔs te*　*trabay,*　*dəpəy yo vẽy*　*i*　*lɔg*　*yo*
　　　　saturday 1p have.NPST work then 1s come.NPST and immediately 1s
　　　　vay　*pə mis nə?*
　　　　go.NPST to mass
　　　　'Saturday we have a shift, then I return and I immediately go to mass, you see?'

Such occurrences of directional *pə*-PPs are very rare and probably acrolectal. In DIP, motion verbs normally command either a bare NP or a *nə*-PP.

Very rarely, a preposition *a* can also precede nominal indications of time, such as in *a noyt* 'at night'. However, the most common prepositional indicator of location, either in spatial or temporal terms, is *nə* (sometimes *nu*) - see section 7.4.9.

The preposition *pə* sometimes contracts with the element which follows, provided that it begins with a vocalic segment. Consider for instance the contracted form of *pə* and *es* 'this' > *pes* 'to this'.

7.4.2 *Atɛ*

PPs introduced by both *(a)tɛ* or *ata* 'until' indicate the endpoint of a temporal/spatial stretch such that this temporal/spatial limit coincides with the endpoint of a given predication:

(170)　a.　ten, ten thirty *atɛ* twelve o'clock.
　　　　　　ten ten thirty until twelve o'clock
　　　　　　'From ten, ten thirty until twelve o'clock.'

　　　　b.　*el fuji-w*　*i*　*foy*　*atɛ ikɔl*　line.
　　　　　　3s flee-PST and go.PST until DEM$_d$ line
　　　　　　'He ran and went until the (finishing) line.'

182

If the PP is occupied by an embedded clause none of which arguments is coreferential with any in the matrix clause, the preposition is followed by the relativiser *ki* 'that', as in (171):

(171) *atɛ   ki   nã   cheg-a   ɔr,   dex fik-a.*
       until CMP NEG$_{cl}$ arrive-INF hour let stay-INF
       'Until the time comes, let it be.'

7.4.3   *Də*

The most frequent preposition in DIP, and arguably the one with the widest functional scope, is *də*. Insofar as the NPs it introduces often function as adverbials of origin (172a), one can recognise a generic ablative semantics. Additionally, the preposition also functions as a generic relator allowing a non-prototypical noun modifier - such as a noun or NP - to modify the head of an NP; in this capacity, *də* introduces the possessor in a possessive construction (172b) or a simple qualifier (172c,d).

(172) a. *es   vey   / ɛ   də Gogla.*
           DEM$_p$ come.PST   COP$_i$.NPST *də* Goghla
           'These are/came from Goghla.'

       b. *yo vay   nə   kazəmêt də mĩ   irmã.*
           1s go.NPST LOC wedding *də* 1s.POSS sister
           'I am going to my sister's wedding.'

       c. *liŋ   də Diw.*
           tongue *də* Diu
           'The language of Diu.'

       d. *jə   ɛ   tẽp də fri   təme, nə?*
           already COP$_i$.NPST time *də* cold REQ
           'It's already the cold season, isn't it?'

This preposition is also the essential building block of complex prepositions such as *(i)sim də* 'on top' or *dẽt də* 'inside' even if, as discussed in 6.7, it is occasionally left out from these compounds. For examples, see 7.4.8 and 7.4.10 below.

As described in 7.11.1, standards of comparison in comparative constructions are obligatorily expressed by *də*-PPs. The sentence in (173) is an example of such a construction:

(173) *galiŋ ɛ   mays barat ki   də karner.*
       chicken COP$_i$.NPST more cheap COMP *də* mutton

'Chicken is cheaper than mutton.'

Də-PPs often constitute temporal adverbials of the type exemplified in (174a,b) - see also 7.2.3. In the speech of one young man at least, similar PPs can also indicate physical location (174c), although this is extremely rare:

(174)   a.   *oj*   *də maŋã*   *yo ko*   *Fabian t-iŋ*   *gi-a*   *saykəl.*
           today *də* morning 1s COM F.   IPFV-PST ride-INF bicycle
           'Today, in the morning, me and Fabian rode our bicycles.'

        b.   *d-ikəl*   *vez*   *yo apiŋ-o*   *nə*   *pə*   uncle?
           də-DEM$_d$ occasion 1s meet-PST REQ DAT uncle
           'The other day I met uncle, didn't I?'

        c.   *mure-w*   *də mar.*
           die-PST *də* sea
           '(He) died at sea.'

Də often contracts with subsequent vowel-initial words. This explains, for example, the occurrence of *des* 'of this' as the contraction of *də* and *es* 'this', and also *dikəl* in (174b) as the contraction of *də* and *ikəl*.

While *də*-PPs modifying an NP typically occur postnominally, these constructions are particularly prone to inversion; the PP often occurs before the noun, in which case the preposition may optionally be dropped (175a,b). This tendency affects certain complex prepositions which contain a *də*-PP, as shown in (175c):

(175)   a.   *ɛl tə*   *fik-a*   *med də tud koyz* / *(də) tud koyz med.*
           3f IPFV.NPST become-INF fear *də* all thing   *də*   all thing fear
           'She's always scared of everything.'

        b.   *kaz*   *də tɛtɛ* / *(də) tɛtɛ kaz*   *jə*   *bɛ̃ze-w?*
           house of T.   of   T.   house already bless-PST
           'Has [he] already blessed Tɛtɛ's house?'

        c.   *jũt*   *də mĩ*   / *(də) mĩ*   *jũt*   *nã*   *te*   *diŋer.*
           together of 1s.OBL   of   1s.OBL together NEG$_{cl}$ have.NPST money
           'I have no money.'

The structure of complex prepositions is described in detail in 7.4.8 and 7.4.10.

7.4.4 $\tilde{E}$

The preposition $\tilde{e}$ 'in, at' is quite rare and occurs only in the speech of some of the least basilectal speakers. It is therefore an acrolectal feature, taking on board the equivalent SP preposition *em* 'in, at'. It is probably part of a special crystallised construction, given that all the recorded instances of this preposition involve the referent *kaz* 'house': $\tilde{e}$ *kaz* 'at home'.

7.4.5 *Kawz də*

The preposition *kawz (də)* or *pukawz (də)* 'because of' introduces a (non-clausal) adverbial of reason:

(176) *use vey   aki pukawz də es   purtəgez.*
       2   come.PST here because of DEM$_p$ portuguese
       'You came here because of this Portuguese.'

Like other prepositions, it can occur in an apparent subordinating function, if the clause under its scope is introduced by the relative pronoun *ki*:

(177) *nɔs nã   foy   fɛs   kawz   ki   nɔs t-iŋ   bastãt trabay.*
       1p NEG$_{cl}$ go.PST party because CMP 1p have-PST much work
       'We did not go to the party because [Lit. 'for the reason that'] we had a lot of work (to do).'

In this case, the complex *kawz ki* could equally well be replaced with the subordinator *purke* 'because'- see section 7.10.2.

7.4.6 *Kom/kufɔr*

Two different prepositions mark the standard of similarity in similative constructions. The first one is *kom* 'like'; the second one, *kufɔr* 'the same way as' seems to be restricted to the expression of similarity in manner. Both elements usually introduce finite clauses rather than nominals, and they also occur as relative pro-forms, but the following sentences exemplify their prepositional status:

(178) a. *galiŋ   lɔg   fik,   galiŋ   ɛ   tẽr   kom pex.*
           chicken immediately become.NPST chicken COP$_i$.NPST tender like fish
           'Chicken will get (cooked) quickly, chicken is as tender as fish.'

b. *yo vay kaz brĩk-a mĩ saykəl kufər* dhoom.
 1s go.INF house play-INF 1s.POSS bicycle like   dhoom
 'I am going home to play INS my bicycle like they do in *Dhoom*.'

See section 7.10.2 concerning manner adverbial clauses, and 7.11.2 for a full account of similative constructions.

7.4.7 *Ku/sẽ*

Ku or *ko* 'with' is both an instrumental (179a) and a comitative marker (179b), while the opposite is conveyed by the preposition *sẽ* 'without' (179c):

(179) a. *nɔs tə kum-e ku mãw.*
    1p IPFV.NPST eat-INF INS hand
    'We eat with our hands.'

   b. *yo ku də mĩ amig, nɔs foy Una.*
     1s COM of 1s.OBL friend 1p go.PST U.
     'I went to Una with my friends.'

   c. *sɛrt jẽt pəd abr-i kok sẽ fak.*
     some people can.NPST open-INF coconut without knife
     'Some people can open coconuts without a knife.'

These PPs often express adverbial notions of manner, and they can integrate an AdvP as a full-fledged adverbial head. For an example of this, see (164) in section 7.3.

In addition to its prepositional functions, *ku* is also an NP-coordinator, which is described in section 7.9.1 below.

7.4.8 *Jũt də*

The complex preposition *jũt də* 'along with' is essentially a comitative marker, as demonstrated by the following examples:

(180) a. *rapaz də Gilbert t-iŋ brĩk-a jũt də mĩ kriãs.*
    boy of G. IPFV-PST play-INF together of 1s.POSS child
    'Gilbert's son used to play with my children.'

   b. *yo gɔs kõt-a istər jũt də Conchita.*
     1s like.NPST tell-INF story together of C.
     'I like chatting to Conchita.'

This preposition has the peculiarity of freely allowing the inversion of its components, so that the *də*-marked PP may occur before the element *jŭt*.[23] This is illustrated in (181):

(181) *yo fik   d-ɛl jŭt.*
       1s dwell.NPST of-3f together
       'I live with her.' / 'I live next to her.'

The competing translations given for (181) betray the fact that *jŭt də* also licenses a locative reading (see section 9.5.3). Other examples point more clearly towards a locative interpretation, such as the one in (182):

(182) *jŭt   d-use cher   mal.*
       together of-2   smell.NPST badly
       'It smells bad near you.'

In addition, *jŭt də* normally marks certain subjects of the possessive verb *te* 'to have' in order to indicate temporary possession (see also 9.5.2):

(183) *pərki   d-ɛl jŭt   nã   t-iŋ   diŋer.*
       because of-3f together NEG$_{cl}$ EXS-PST money
       'Because she had no money.' [Lit. 'There was no money next to her.']

With respect to the alignment of possessive subjects, see also section 7.5.

7.4.9   *Nə*

The standard locative preposition is *nə* (sometimes also *nu*). It must be pointed out that, with motion and locative verbs, locative phrases most of the times involve no preposition at all, although they optionally admit them. This is illustrated in the following example:

(184) *el foy   / t-iŋ   (nə) Diw.*
       3s go.PST   EXS-PST   LOC Diu
       'He went to / was in Diu.'

Whenever location does not derive from the inherent semantics of the verb, *nə* is employed. Consider the examples in (185):

---

[23]The exceptionality of this particular complex preposition with respect to word order is not circumscribed to Diu IP; a similar possibility is reported for both Daman IP and Korlai IP by Clements and Koontz-Garboden (2002:223).

(185) a. *pərkə   aki   tir-o*   shooting *nə   Diw.*
           because here take-PST shooting LOC Diu
           'Because (they) shot (it) here in Diu.'

       b. *dəpəy atər-o   pɛd   nə   ag.*
           then   throw-PST stone LOC water
           'Then he threw stones into the water.'

       c. *nə   saykəl* uncle *t-iŋ   vay.*
           LOC bicycle uncle IPFV-PST go.INF
           'Uncle was going by bicycle.'

One will notice that, in all of the above, *nə* signals generic location only, without any specification of relative position. Further locative specification may be achieved with the complex prepositions discussed in section 7.4.10.

The type of location construed with *nə* may be temporal as well as spatial. The following is an example of the temporal use of this preposition:

(186) *vey   nə*   two thousand one *vey.*
       come.PST LOC two thousand one come.PST
       'I came in two thousand and one.'

The preposition may optionally contract with the following word if it is vowel-initial, in which case e.g. *nə + ikəl di* 'that day' > *nikəl di* 'on that day'.

7.4.10   Other complex prepositions

Other than *jũt də* and *(pu)kawz də* 'because of' (separately discussed in sections 7.4.5 and 7.4.8), all complex prepositions have clear locative (temporal as well as spatial) semantics and therefore primarily introduce adjuncts. Those include *ãt də* 'before', *bayx də* 'under', *dəpəy də* 'after', *dẽt də* 'inside', *durãt də* 'during', *fɔr də* 'outside', *frẽt də* 'in front of', *lad də* 'next to', *lõj də* 'far from', *pɛrt də* 'close to', *(i)sim də* 'on top of' and *tras də* 'behind'. *Jũt də*, as mentioned in 7.4.8, does license a locative reading, although this is not the single defining trait of the preposition.

Some examples of complex prepositions in use with (temporal) locative functions follow:

(187) a. *kar   tud tə   rãk-a   də panɛl tə   por   dẽt*
           meat all IPFV.NPST remove-INF from pan   IPFV.NPST put.INF inside
           də jel.
           of ice
           '(I) remove the meat and all that from the pan and put it in ice.'

b. *dəpəy də trey di use vẽy volta-d.*
after of three day 2 come.NPST return-PTCP
'Come back in three days.'

c. *sɔ durãt də nɔven.*
only during of novena
'Only during novenas.'

d. *ãt də mat-a buf nã vẽd pɛl.*
before of kill-INF buffalo NEG$_{cl}$ sell.NPST skin
'Do not sell the skin before killing the buffalo.'

e. *ikəl muyɛr fik də mĩ lad.*
DEM$_d$ woman dwell.NPST of 1s.OBL side
'That woman lives next to me.'

The examples in (187) reveal some interesting characteristics of the complex prepositions. The various PPs here contain not just bare nouns (187a,c), but also fuller NPs (187b), an oblique personal pronoun (187e) and a nonfinite clause (187d). Finite clauses, on the other hand, are disallowed from this syntactic environment; they may occur with a very similar function, however, if they are made to include the subordinator *ki* 'that'.[24] A finite equivalent to the nonfinite clause in (187d), a local saying, would be as follows:

(188) *ãt ki use mat buf nã vẽd pɛl.*
before CMP 2 kill.NPST buffalo NEG$_{cl}$ sell.NPST skin
'Do not sell the skin before you kill the buffalo.'

One final observation is that, as shown also for *jŭt də*, the lexical component of complex prepositions (e.g. *durãt*, *dəpəy*, etc.) may in fact occur in final position. This is exemplified in (187e). It appears, however, that none of the complex prepositions described in this section do so as frequently as *jŭt də* (see 7.4.8).

7.5 Argument alignment

Argument alignment refers to the system governing the morphosyntactic characteristics - case marking, for instance - of arguments across different constructions. Research on alignment strategies other than nominative/accusative has lately deconstructed the crosslinguistic validity of the notions of Subject and Object, and novel descriptive terminology has been proposed (see Comrie 1978, Dixon 1979, Croft 2001, Hengeveld and Mackenzie 2008, Malchukov et al. forthcoming). In order not to enforce any preconceived alignment model

---

[24]This possibility applies to all complex prepositions with temporal semantics.

onto DIP, and considering that Gujarati - one of the languages in the equation of DIP's formation and development - displays split ergativity (more information in 9), I will follow the proposals of such studies in discussing matters of alignment and word order (see also section 7.6). Arguments will accordingly be labelled as follows:

- The sole argument of an intransitive clause is indicated with "S";

- In monotransitive clauses, "A" stands for the most agent-like argument, while "P" stands for the least agentive element;

- In ditransitive clauses, "A" stands for the agentive argument, "U"[25] stands for the undergoer and "R" indicates the recipient argument.

Whereas purely syntactic functional categories (such as subject or object) have been paramount in traditional linguistic descriptions, it is now clear that alignment can reflect different properties of the arguments, such as their semantic and pragmatic status. Hengeveld and Mackenzie (2008:317) distinguish three major types of alignment, from a crosslinguistic perspective, admitting that a particular language may exhibit a combination of elements from the various domains. The first of these, *interpersonal alignment*, refers to an organisation of arguments which is sensitive to their pragmatic status (e.g. whether or not they receive focus) or referential properties (e.g. specificity). Languages exhibiting *representational alignment* are those whose arguments are marked according to their semantic functions (e.g. recipient, experiencer) or properties (e.g. animacy). Finally, *morphosyntactic alignment* refers to the type of alignment which takes into account the syntactic functions of the arguments (e.g. subject or object) or morphosyntactic measures of complexity and weight.

As described below, it is clear that the *representational* level is responsible for a great deal of the alignment strategies of DIP. The morphosyntactic properties of arguments in DIP depend primarily on their semantic function, to the extent that it is convenient to divide the present discussion into two separate sections dealing with the case-marking properties of animate versus inanimate arguments respectively. Across languages, notions of animacy, definiteness and person often interact with alignment strategies, and this is also the case in DIP. I must therefore introduce the *individuation scale*, an implicational hierarchy conflating animacy, definiteness and person scales. Haspelmath (2007a:84) finds that ditransitive alignment splits (i.e., alignment strategies displaying differential marking for different types of referents in similar argument position) usually occur at a given step of this scale, which he organises as follows:

1st/2nd > 3rd > proper noun > human > non-human

---

[25] The notation of this argument varies in the literature. Whereas I opt for "U" in accordance with for instance Hengeveld and Mackenzie (2008), some authors use "T", for "Theme" (e.g. Haspelmath 2007a, Malchukov et al. forthcoming). It is important to be aware of the interchangeability of these labels when approaching the literature.

This scale will be recuperated in the following sections as explanation for specific patterns of case-marking observed in DIP.

Alignment types are manifested in various ways in different languages, including case-marking, the use of adpositions, agreement and so forth. In DIP, the function of arguments is flagged through the use (or absence) of adpositions (see section 7.4). Some prepositions which are particularly relevant for the discussion at hand include the dative cluster $a/pə$ (described in 7.4.1) as well as the locative *jũt də* 'near, next to' (see 7.4.8) - and occasionally also the comitative *ku*, a functional extension which is possibly due to the fact that *ku* and *jũt də* are equivalent in their role as comitative markers (see sections 7.4.7 and 7.4.8, also this chapter fn. 12).

7.5.1 Alignment of inanimate arguments

An inanimate S argument is left bare, i.e., unmarked by any preposition. This is the case in example (189):

(189)   *cher tə   vi.*
        smell IPFV.NPST come.INF
        'It smells [Lit. 'Smell comes'].'

In monotransitive clauses, inanimate A's and P's are prototypically bare. Consider (190), in which both arguments are inanimate:

(190)   *aviãw arəm-o   pared.*
        plane hit-PST wall
        'The plane hit the wall.'

Inanimate P's can optionally be Dative-marked, which is particularly common if they are perceived as a target or a beneficiary in some sense. A construction in which a P is Dative-marked is usually semantically distinct from a similar construction with the same verb and a bare P. Consider the difference between the following two sentences, both of which involve the lexical verb *ve* 'to see':

(191)   a.   *yo vay   ve   sinem.*
             1s go.NPST see.INF film
             'I am going to watch a movie.'

        b.   *vay   ve   pə   leyt.*
             go.NPST see.INF DAT milk
             'Go check on the milk.'

The sentence in (192) further exemplifies the possibility to Dative-mark inanimate P's.

(192)   *vẽt   apəg-o   pə   vɛl.*
        wind blow.out-PST DAT candle
        'The wind blew out the candle.'

The *individuation scale* introduced in 7.5 now gains relevance when we approach inanimate (or animate, for that matter) pronominal P's, which are almost always dative marked. Recall that the highest positions of the scale are occupied by pronominal forms, and in fact the higher the position in the individuation scale, the most likely it is that a P argument will be dative marked in DIP - it is important to stress that this tendency is valid for inanimate P's but not for inanimate S's or A's, which indicates a tendential nominative/accusative alignment for DIP. Two examples are given in (193). Although the pronoun refers to a type of meat in (193a), the fact that *karner* is an animal may raise doubts as to the inanimacy of the argument; in (193b), however, the referent is clearly inanimate, as it stands for a mix of ingredients.

(193)   a.   *karner   ɛ   dur   nə,   ti   ki   kuz-e   bẽy a   el.*
             mutton COP$_i$.NPST hard REQ AUX.NPST CMP cook-INF well DAT 3s
             'Mutton is hard, it has to cook well.'

      b.   *sɔ   mw-er   a   el, dəpəy põy   sum də limãw.*
             only grind-INF DAT 3s then   put.NPST juice of lime
             '[You] only [have to] grind it, and then add a little lime juice.'

In ditransitive constructions, U's are bare, as shown for *prɛsẽt* in (194a). Because ditransitive constructions typically involve transfer of possession, inanimate A's and R's are extremely rare, and when they do occur they can often be interpreted as animate entities. Such is the case of agent *Go* 'Goa' in (194b) and recipient *igrej* 'church' in (194c), which is Dative-marked:

(194)   a.   *yo mãd-o   ũ   prɛsẽt a   mĩ   irmã.*
             1s send-PST one present DAT 1s.POSS sister
             'I sent a present to my sister.'

      b.   *Go   nə   ad   da   rɛpos.*
             Goa NEG$_{cl}$ IRR.NPST give.INF answer
             'Goa will not give [me] an answer.'

      c.   *nɔs de-w   / mãd-o   diŋer pə   igrej.*
             1p give-PST   send-PST money DAT church
             'We gave/sent money to the church.'

In sum, if one takes into account inanimate arguments only, the following picture obtains:

1) all arguments are typically bare except for R;
2) R is obligatorily Dative-marked; and
3) P (but not U) is optionally Dative-marked.

This is captured in Figure 7.1. The continuous line indicates the basic alignment of inanimate arguments, whereas the dotted line represents an optional (secondary) alignment:
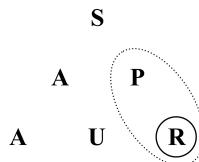
Figure 7.1: Alignment of inanimate arguments in DIP

What we have, therefore, is a system in which arguments of intransitive and monotransitive clauses are either non-distinct or display a nominative-accusative alignment; as far as the relationship between monotransitive and ditransitive clauses is concerned, the system is potentially of the primative-secundative type (see Hengeveld and Mackenzie 2008, Malchukov et al. forthcoming).

7.5.2 Alignment of animate arguments

Animate arguments partly reflect the alignment of inanimate arguments described in the previous section, but with more solid and widespread patterns of dative-marking.

Animate S arguments, whether or not pronominal, tend to be bare, as in (195):

(195)   a.   officer *nə*   *ad*   *gost-a*.
                officer NEG$_{cl}$ IRR.NPST like-INF
                'The officer won't like [it].'

          b.   *yo nã̃   sab*.
                1s NEG$_{cl}$ know.NPST
                'I don't know.'

In the case of monotransitive constructions, on the other hand, the case assigned to animate A arguments is dependent upon their semantic function. Agents are bare even if pronominal, as in (196):

(196) a. *Armando tə   fal-a   mĩtir.*
           A.   IPFV.NPST speak-INF lie
           'Armando is telling lies.'

       b. *el mem atɛr-o   a   el.*
           3s EMPH push-PST DAT 3s
           'He pushed him.'

If A is construed as the experiencer of an action beyond their control, however, it is Dative marked, as in (197):

(197) a. *a   mĩ   tə   sĩt-i   fri.*
           DAT 1s.OBL IPFV.NPST feel-INF cold
           'I'm cold.' [Lit. 'To me is feeling cold.']

       b. *a   mĩ   nã   apaŋ   tẽp.*
           DAT 1s.OBL NEG_cl find.NPST time
           'I don't find the time.'

       c. *pə   ɛl təme aprende-w purtɛgez   ku   nɔs.*
           DAT 3s also learn-PST Portuguese COM 1p
           'She also learnt Portuguese from us.'

       d. *pə   el fik-o   bẽ duẽt nə.*
           DAT 3s become-PST very sick REQ
           'He fell really sick, you see.'

Verbs of possession are particularly interesting, as there is much variation concerning the A argument. A's may be bare, as in (198a), or they may be Dative-marked, as in (198b,c). Additionally, there is a construction in which the A is marked with the locative/comitative preposition *jŭt də*, and occasionally comitative *ku*, indicating temporary possession; this is illustrated in (198d,e), see also 7.4.8:

(198) a. *yo nã   te   nĩge   kaz.*
           1s NEG_cl have.NPST nobody home
           'I do not have anyone at home.'

194

b. *te   bigɔd   pə   el.*
have.NPST moustache DAT 3s
'He has a moustache.'

c. *a   el te   med.*
DAT 3s have.NPST fear
'He is scared [Lit. 'He has fear'].'

d. *mĩ   jŭt   nã   te   muyt diŋer nã   te.*
1s.OBL together NEG$_{cl}$ have.NPST much money NEG$_{cl}$ EXS.NPST
'I don't have much money.' [Lit. 'There isn't much money near me.']

e. *ku   el nə   t-iŋ.*
COM 3s NEG$_{cl}$ have-PST
'He did not have [it].'

Animate P arguments are Dative-marked, as shown in the following examples - see also (196b) above:

(199)   a. leopard *foy   murd-e pə   lion.*
leopard go.PST bite-INF DAT lion
'The leopard went and bit the lion.'

b. *vay   cham-a pə   Bablu.*
go.INF call-INF DAT B.
'Go call Bablu.'

c. *el kɛr   kum-e a   mĩ   nə   istrad.*
3s want.NPST eat-INF DAT 1s.OBL LOC road
'He wants to eat me on the road.'

As in the case of the inanimate arguments, ditransitive constructions containing animate arguments reveal an interesting characteristic of DIP. Unlike P in monotransitive templates, animate U in ditransitive constructions is typically bare. The following two sentences reveal the different treatment of the same animate non-human argument *kãw* 'dog' as a P (200a), in which it is dative-marked, and as a U (200b), in which it is left bare:

(200)   a. *Virgilo bate-w   pə   kãw.*
V.   hit-PST DAT dog
'Virgilo hit the dog.'

b. *Virgilo de-w   kãw pə   Severina.*
   V.   give-PST dog DAT S.
   'Virgilo gave the/a dog to Severina.'

The absence of dative-marking on U is motivated by the dative-marking of the R argument - such as *Severina* in (200b) -, in a classic primative-secundative alignment. R is obligatorily dative-marked in DIP, as shown in (201):

(201) *yo a   da   kriãs pə   tɛtɛ.*
       1s IRR.NPST give.INF child DAT aunt
       'I will give the children to [my] aunt.'

The general primative-secundative strategy however clashes with a tendency to dative-mark arguments (other than S and A) high in the individuation scale - see section 7.5. There is in effect a split in the assignment of case to U which allows dative-marking on pronominal arguments and proper nouns only. Some examples follow; (202b) is a transformation of (201) above:

(202) a. *yo mãd-o   / lɛv-o   pə Liza pə   pray.*
          1s send-PST   take-PST DAT L.   DAT beach
          'I sent/took Liza to the beach.'

       b. *yo a   da   a elz pə   tɛtɛ.*
          1s IRR.NPST give.INF DAT 3p DAT aunt
          'I will give them to [my] aunt.'

It is interesting to notice that demonstratives do not follow this trend, i.e., they are not treated the same way as personal pronouns. In the corpus, any U occupied by a demonstrative is bare:

(203) *use da   es   pə   mãy.*
       2s give.NPST DEM$_p$ DAT mother
       'You, give this to mother.'

With respect to animate arguments, then, we observe the following:

1) S is unmarked;
2) A is optionally case-marked;
3) P is Dative-marked;
4) U is unmarked, except if pronominal or proper noun; and
5) R is obligatorily Dative-marked;

Figure 7.2 is a representation of these facts:
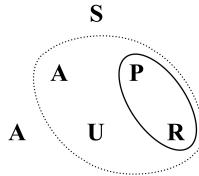
Figure 7.2: Alignment of animate arguments in DIP

Animate arguments therefore constitute a clearer manifestation of the primative-secundative alignment - i.e. the tendency to case-mark P and R similarly and U differently - which is, however, disturbed by a tendency to assign dative case to non-S/A arguments high up in the individuation scale. In addition, non-bare occurrences of A confirm that case-marking in DIP is highly dependent on semantic considerations. Given that S seems to be consistently bare, one may not discard the syntactic category of Subject in DIP; however, the fact that A's can be case-marked means the subject category is semantically overruled.

7.6 Clausal structure

This section describes the structure of DIP clauses, focusing in particular on the issue of word order. In the interest of cross-linguistic comparability, the arguments are labelled as explained in 7.5 above, while verbs are indicated with the label "V".

One preliminary remark to this descriptive section is that DIP is characterised by widespread ellipsis as well as frequent constituent doubling. Despite these phenomena - explored in sections 7.8.2 and 7.8.3 respectively -, it is possible to ascertain the language's *prototypical* word order(s) if one factors in only cases in which a verb's valency is fully expressed. One should however be forewarned that real speech looks very different from these idealised syntactic templates. It will also be made clear, in section 7.8, that neither ellipsis nor constituent doubling occur randomly, which adds interesting new dimensions to the study of DIP syntax.

7.6.1 Simple declaratives

DIP word order is sensitive to pragmatic considerations, as we shall see below. The basic word order, i.e. the default word order in an affirmative sentence when all information is equally new and no particular focus is to be achieved, is S-V in intransitive clauses, A-V-P in monotransitive clauses and A-V-U-R in ditransitive clauses.[26] In traditional terms, then,

---

[26]See also the examples provided in section 7.5.

DIP would fall under the S-V-O category. The first example under (204) was elicited[27], the following sentences occurred in unconstrained speech. All of these exemplify the S-V structure of intransitive clauses. (204a,b) show that animacy considerations do not enter the definition of the position of S in relation to the V. The sentence in (204c) is an instance of the existential use of the verb *te*, which is intransitive in this function, as described in detail in section 7.1.3.3:

(204) a. *leyt vaz-o.*
           milk spill-PST
           'The milk spilt.'

       b. *Armando kai-w.*
           A.   fall-PST
           'Armando fell down.'

       c. *ŭ   makak t-iŋ.*
           one monkey EXS-PST
           '[Once upon a time,] there was a monkey.'

The example sentences in (205) exemplify the basic A-V-P word order of transitive clauses:

(205) a. *ɔm larg-o   ped.*
           man release-PST fart
           'The man let out a fart.'

       b. *yo tə   kuziŋ-a   aroz ku pex.*
           1s IPFV.NPST cook-INF rice and fish
           'I am cooking rice and fish.'

The basic word order of ditransitive constructions is A-V-U-R, regardless of the animacy value of the various arguments. The sentences in (206) exemplify this pattern:

(206) a. *nɔs de-w   diŋer pə   igrej.*
           1p give-PST money DAT church
           'We gave money to the church.'

---

[27]Despite its pragmatic shortcomings, data obtained as elicited translation is particularly useful to ascertain basic word order precisely because the task minimises contextual interference and ensures that all the information is equally new.

b. *Virgilo de-w   kãw pə   Severina.*
   V.   give-PST dog DAT S.

   'Virgilo gave the/a dog to Severina.'

Departures from these basic structures constitute pragmatically marked constructions which react to particular pragmatic conditions. They either place particular emphasis on P and U, or target the occupation of certain syntactic positions left empty by ellipsis. For a full account of the syntactic effects of pragmatic conditions, see section 7.8.

7.6.2   Interrogatives

Like many other languages, DIP allows for three types of interrogative clauses: the so-called *polar questions* (also called *yes/no questions*, those which elicit confirmation or disconfirmation of a given state of affairs), *tag-questions* (which can be seen as a subset of polar questions, but syntactically and functionally peculiar) and *content questions* (also called *q-word questions*, questions which target either a particular clausal constituent or an entire sentence and usually warrant a complex reply). Word order is more prominent in the formation of content questions than polar interrogatives. I will discuss these in turn.

7.6.2.1   Polar questions

Polar questions are those typically used for the elicitation of unknown information or the (dis)confirmation of of an expectation, i.e. whose answer is expected to be either 'yes' or 'no'. In DIP, polar questions do not involve any peculiar syntactic construction. They are marked by prosody alone, which consists of applying a rising intonation towards the end of the sentence.

(207) a. *ali   ikəl   gate nã   te   apərs-e   pikənin?*
           there DEM$_d$ gate NEG$_{cl}$ IPFV.NPST show-INF small

           'Do you see that small gate over there?' [Lit. 'Doesn't that small gate show over there?']

       b. use tə fik-a med də lagərtix?
           2 IPFV.NPST become-INF fear of gecko
           'Are you scared of geckos?'

The presence of clausal negation in a polar question, as in (207a), indicates a counter-expectation, i.e., the expectation that the veracity of the polar opposite will be confirmed.

7.6.2.2 Tag questions

Tag questions make use of a tag - the requestative particle *nə*[28] - which typically occurs at the end of the clause (for changes to the prototypical word order, see 7.8). The function of tag questions is not solely to ascertain unknown information, and even when they are they entail the expectation of a positive assertion:

(208)   a.   *aki   nə   Diw nɔs fal   asĩ* biting, *nə   Savid?*
             here LOC Diu 1p   say.NPST thus *biting*   REQ S.
             'Here in Diu we say *biting*, don't we Savid?'

         b.   *el tə   beb-e   sɔ   leyt nə?*
             3s IPFV.NPST drink-INF only milk REQ
             'He drinks nothing but milk, doesn't he?'

         c.   *jə   ε   tẽp   də fri   təme, nə?*
             already COP_i.NPST time of cold also   REQ
             'It's also the time of cold already, isn't it?'

In (208b,c), the information being questioned may or may not be unknown to the speaker. This brings us to the primary function of tag questions; tags are commonly pragmatic means to establish the clarity of an assertion/speech act (209a,b) or elicit the agreement/acknowledgement of the interlocutor (209c). This is particularly evident in the following examples. The elicitation of unknown information is clearly not relevant in all of these examples, as all assertions refer to the speakers themselves; the speakers are therefore in a privileged position to establish their veracity, and these questions are often rhetorical.

(209)   a.   *nɔs t-iŋ   isim   nə?*
             1p   EXS-NPST above REQ
             'We were upstairs, you see?'

         b.   *yo nə   te   kaz   nə   te   nə?*
             1s NEG_{cl} have.NPST house NEG_{cl} have.NPST REQ
             'I do not have a house, you see?'

         c.   *mĩ   nom   ε   muyt kõprid nə?*
             1s.POSS name COP_i.NPST very   long   REQ
             'My name is very long, don't you think?'

---

[28]This is only one of the functions of the *requestative* element *nə*, which also operates in imperative clauses (see 7.6.3.1). In the sense that a tag question has the purpose of eliciting agreement or acknowledgement, the similarity with a requestative particle is obvious. The tag is accordingly glossed *REQ*, for consistency.

The following sentences exemplify alternative syntactic placements of the tag/requestative particle, viz. immediately after the verb (210a) plus close-knit verbal arguments (210b):

(210)   a.   *d-ikəl*   *vez*   *yo apiŋ-o*   *nə*   *pə*   uncle*?*
             of-DEM$_d$ occasion 1s meet-PST REQ DAT uncle
             'Once I met uncle, wasn't it?'

        b.   *Liza õt*   *nã*   *t-iŋ*   *med nə*   *pə*   uncle*?*
             L.   yesterday NEG$_{cl}$ have-PST fear REQ DAT uncle
             'Yesterday, Liza wasn't afraid of uncle, right?'

In such cases, *nə* does not seem to signal a clausal boundary. If that were the case, one would have to interpret the constituents which occur after *nə* as appended to the clause - cf. the discussion in 7.8.2.

7.6.2.3   Content questions

In content question formation, DIP makes use of a series of interrogative pro-forms (see 6.6.3) apart from adnominal interrogatives (see 6.5.5). The latter are not dealt with in this section, as they have already been dealt with in section 7.2.4, but I exemplify their use in (211) for clarity:

(211)   a.   *use nə*   *kwɔl* *kam ad*   *durm-i?*
             2   LOC which bed IRR.NPST sleep-INF
             'Which bed will you sleep in?'

        c.   *use kwõt*   *an*   *te?*
             2   how.much year have.NPST
             'How old are you?' [Lit. 'How many years do you have?']

        d.   *use ki*   *sirvis tə*   *faz-e?*
             2   what job   IPFV.NPST make-INF
             'What service do you do?'

The elicitation of properties/qualification is often operated with resort to an adnominal interrogative construction involving a generic noun with the meaning of 'type' or 'kind'. This is exemplified in (212):

(212)   *es*   *rɛlɔj* *kwal* *kwalidad də rɛlɔj* *ɛ?*
        DEM$_p$ watch what type   of watch COP$_i$.NPST
        'What type of watch is this watch?'

Adnominal interrogatives, by definition, occur within the NP. Notice that, in (212), *kwal kwalidad də rɛlɔj* functions as an interrogative element in block, the reason why the entire NP occurs in (pre-verbal) focus position. Interrogative pro-forms, on the other hand, replace entire constituents of the clause. *Kom* 'how/why' and *pərki/purki* 'why' interrogate manner and reason adverbials, while *kwɔn* 'when' and *ɔn* 'where' replace time and locative adverbials respectively. *(U)ki* 'what' and *kẽ* 'who' are interrogative pronouns which typically stand for non-human/inanimate and human/animate arguments respectively.

Interrogative pro-forms always occur in the position immediately preceding the VP (i.e., the verbal form plus any TAM markers), which is a focus position in DIP - see 9.6.1 for a comparison with Gujarati. The sentences in (213) exemplify each one of these interrogative pro-forms.

(213)   a.   *use õt   kom nã   vey?*
             2s   yesterday why NEG$_{cl}$ come.PST
             'How come you didn't come yesterday?'

        b.   *use purki tə   da   salt?*
             2s   why   IPFV.NPST give.INF jump
             'Why are you jumping?'

        c.   *use kwɔn a   vay?*
             2   when IRR.NPST go.INF
             'When will you leave?'

        d.   *use ɔn   tə   fik-a?*
             2s   where IPFV.NPST stay-INF
             'Where do you live?'

        e.   *use uki   kɛr?*
             2s   what want.NPST
             'What do you want?'

        f.   *kẽ de-w   es   a   use?*
             who give-PST DEM$_p$ DAT 2
             'Who gave you that?'

Interrogative pro-forms can be modified by prepositions, in accordance with the functions of the constituents they replace. This is demonstrated in (214):

(214)   a.   *es   də kẽ ɛ?*
             DEM$_p$ of who COP$_i$.NPST
             'Whose is this?'

b. *es   pə   kẽ   ɛ?*
    DEM$_p$ DAT who COP$_i$.NPST
    'Who is this for?'

c. *use də ki   tə   fik-a   med?*
    2   of what IPFV.NPST become-INF fear
    'What are you afraid of?'

Apart from the interrogated constituent, the remaining clausal constituents in general occur in the same positions as in non-interrogative clauses, and they are equally affected by pragmatic conditions. For further information, see section 7.8. The syntax of indirect questions is described in 7.10.1.

7.6.3   Imperatives

This section explores the syntax of clauses which intend to command a participant of the speech act to act in a particular way. I will divide the discussion into three distinct parts, dealing in turn with *imperatives*, *prohibitives* and *hortatives*.

7.6.3.1   Imperative

In an imperative construction, the speaker typically directs a given command or request at a second person (addressee). In general, there is no specific imperative verb form in DIP.[29] In this type of clauses, verbs can take two different forms, viz. non-Past (215a) and Infinitive (215b). The latter strategy is slightly more insistent than the former and therefore less respectful:

(215) a. *fal   Liza, fal,   lɔg   fal.*
           speak.NPST L.   speak.NPST immediately speak.NPST
           'Speak up, Liza, speak up, speak up quickly.'

       b. *atər-a   es,   bay.*
           throw-INF DEM$_p$ girl
           'Throw it, girl.'

In (215), the addressees are expressed, but this is often not the case. In case it is clear from context who the command is directed at, there is usually no overt indication, as in (216):

---

[29]The only verb for which a distinct imperative form was recorded is *ver* 'to see', which can employ the imperative form *vej* (from Ptg. *veja* 'see.IMP') in competition with the non-Past form *ve* and the infinitive form *ver*.

(216) a. *say   d-ali.*
           leave.NPST of-there
           'Get out of there.'

      b. *sĕt-a   nə   chãw.*
           sit-INF LOC ground
           'Sit down on the ground.'

It is very common to append the particle *nə* to imperatives; it can occur either immediately after the verb (217a) or at the end of the clause (217b). In either case, the particle is phonetically prominent and intonationally detached from the main clause.

(217) a. *atər-a   nə   fɔr.*
           throw-INF REQ out
           'Throw it away.'

      b. *atər-a   fɔr nə.*
           throw-INF out REQ
           'Throw it away.'

In these contexts, *nə* functions as a requestative particle, resulting in a respectful yet firm command. The examples under (217) show that the particle can be used with Infinitive imperative constructions, but it also occurs with non-Past imperatives, as demonstrated in (218) below:

(218) a. *fal   nə   nə   uvid*
           speak REQ LOC ear
           'Speak into [my] ear.'

      b. *use vay   vist-i   nə.*
           2   go.NPST dress-INF REQ
           'Go get dressed.'

      c. *kalad fik   nə.*
           quiet become.NPST REQ
           'Be quiet.'

Word order in imperatives is rather flexible, in particular given that these constructions tend towards minimal expression and, as a result, ellipsis is rampant. Notice, for instance, how the addressee occurs at the end of the clause in (215b) but not in (218b). In (218c), the order of the verb and the complement is atypical, but this is easily explained with resort to pragmatic considerations: the preverbal position is a focus position in DIP, and by

inverting the canonical word order the speaker is highlighting the most important element of the request (in this case, the intended 'quietness'). For further information, see 7.8 below.

7.6.3.2 Prohibitive

Prohibitives simply involve the negation of the imperative clause - for a detailed account of negation in DIP, see section 7.7. Apart from the fact that verb forms in prohibitives are non-Past rather than infinitive, they are very similar to imperative clauses. Some examples can be found in (219):

(219) a. ãt   də mat-a   buf   nã   vẽd   pɛl.
         before of kill-INF buffalo NEG$_{cl}$ sell.NPST skin
         'Do not sell the skin before killing the buffalo.'

      b. nã   lɛv   mal, jẽt   pɔb.
         NEG$_{cl}$ take.NPST badly people poor
         'Don't feel offended [Lit. 'Don't take [it] badly'], [we are] poor people.'

Requestative *nə* also operates in prohibitives, enhancing the vehemence of the prohibition, as shown in (220):

(220) nã   faz   es   nə.
      NEG$_{cl}$ make.NPST DEM$_p$ REQ
      'Don't do that.'

7.6.3.3 Hortative

In the hortative construction, the hortative particle *bam* 'let's' usually combines with the infinitive form of the main verb to express resolve (221a), encouragement or strong suggestion (221b) of a course of action. The agent-like argument of a prototypical hortative construction is deictic and, most of the times, left unexpressed. Hortative constructions function as first person imperatives, as they are always directed at the speakers themselves and optionally also the addressee. In narratives, this construction is rather common in reported speech (221c).

(221) a. dəpəy də afternoon pẽs-o   bam   faz-e   ũ piken koyz dẽt   də
         after of afternoon think-PST HORT make-INF one small thing inside of
         famil.
         family
         'Then, in the afternoon, [I] thought: "Let's do something small for the family".'

205

b. *bam   fuj-i.*
   HORT flee-INF
   'Let's flee.'

c. *oj   otrez el də* afternoon *tə   diz-e   a   mĩ   ki*
   today again 3s of afternoon IPFV.NPST say-INF DAT 1s.OBL CMP
   bam   faz-e   kwɔlkɛ chaknũ.
   HORT make-INF any   snack
   'Today again they told me: "Let's make some snacks".'

The corpus contains some examples in which *bam* co-occurs with an expressed subject, as in (222):

(222) *bam   nɔs kõt-a   istər jũt   d-es   uncle.*
       HORT 1p   tell-INF story together of-DEM*p* uncle
       'Let's both of us chat with this uncle.'

In the rare occasions in which *bam* does not govern a verb, what follows is a locative phrase and the clause has an allative meaning:[30]

(223) *bam   pray, pasy-a   pray.*
       HORT beach take.walk-INF beach
       'Let's go to the port, take a walk around the port.'

This shows that the particle has some inherent semantics of motion towards a given place. Etymologically, this particle derives from a form of the Portuguese verb 'to go' (*vamos* 'we go/let's go'), but *bam* has none of the formal characteristics of a verb in DIP.

Bam also occurs as an interjection, although less commonly. This is clearly the case in (224); the fact that the main verb is TAM-marked is evidence of the clausal divide between *bam* and the finite clause:

(224) *bam   nɔs doy a   faz-e   race.*
       HORT 1p   two IRR.NPST make-INF race
       'Come, the two of us will make a race.'

7.6.4   Adverb placement

Adverbs, as explained in 6.3, constitute a rather heterogeneous class. It is therefore not surprising to find that different types have different distributional characteristics. In general, the position of adverbs in the clause is rather flexible, but that is particularly so for

---

[30] In (223), the word *pray* 'beach' is translated as 'port' given that this is the DIP name for the port area of Diu Town - see 8.6.4 for further information.

manner, temporal and spatial adverbs. Consider the recorded placement of the manner adverb *lɔg* 'immediately, quickly':

(225)   a.   *galiŋ*   *lɔg*   *fik,*   *galiŋ*   *ε*   *tẽr kom pex.*
             chicken immediately become.NPST chicken COP$_i$.NPST soft like fish
             'Chicken gets done quickly, chicken is soft like fish.'

        b.   *mĩ*   *mãy*   *lɔg*   *cham-o*   *pə*   *dotor.*
             1s.POSS mother immediately call-PST DAT doctor
             'My mother immediately called the doctor.'

        c.   *dəpəy yo vẽy*   *i*   *lɔg*   *yo vay*   *pə*   *mis*   *nə?*
             after   1s come.NPST and immediately 1s go.NPST DAT mass REQ
             'Then I come and I immediately go to mass, you see?'

        d.   *use akab*   *kum-e*   *lɔg*   *i*   *vay*   *imɔr.*
             2   finish.NPST eat-INF immediately and go.NPST away
             'Finish eating quickly and go away.'

        e.   *oj*   *t-iŋ*   *iskɔl*   *may termin-o*   *lɔg.*
             today EXS-PST school but   finish-PST immediately
             'There was school today, but it finished early.'

In the previous sentences, the adverb *lɔg* occurs immediately prior to the verb (225a,b), clause-initially (225c) and clause-finally (225d,e). A rather similar picture obtains for the distribution of the temporal adverb *õt* 'yesterday', as demonstrated in (226):

(226)   a.   *õt*   *nɔs t-iŋ*   *ve*   *teytə.*
             yesterday 1p   IPFV-PST see.INF theater
             'Yesterday we went to the theater.' [Lit. 'Yesterday we saw the theater.']

        b.   *õt*   *t-iŋ*   *ti-d*   *mis.*
             yesterday IPFV-PST EXS-PTCP mass
             'There was a mass yesterday.'

        c.   *Liza õt*   *nã*   *t-iŋ*   *med nə*   *pə*   uncle*?*
             L.   yesterday NEG$_{cl}$ have-PST fear REQ DAT uncle
             'Liza wasn't scared of *uncle* yesterday, right?'

        d.   *use õt*   *uki fez?*
             2   yesterday what do.PST

'What did you do yesterday?'

The adverb *õt* occurs in clause-initial position in (226a,b), immediately preceding the VP in (226c) and before the preverbal focus position (which is occupied by the interrogative pro-form *uki*) in (226d).

The placement of adverbs within imperative sentences is likewise flexible, as demonstrated by the sentences in (227):

(227) a. *Leslie, vay   may tras, may may.*
           L.   go.NPST more back more more
           'Leslie, go further back, further further.'

       b. *muyt* fast *nã   vay,   a   ka-i.*
           very fast NEG$_{cl}$ go.NPST IRR.NPST fall-INF
           'Don't go too fast, you'll fall down.'

Phasal adverbs are much more constrained in their distribution. These elements normally occur immediately before the VP, as shown in (228a,b,c). It is not normal for any clausal constituent to intervene in between a phasal adverb and the VP, but the adverbs sometimes - though rarely - occur in postverbal position (228d):

(228) a. rabbit *ain te   atras.*
           rabbit still EXS.NPST behind
           'The rabbit is still back there.'

       b. *el jə   chig-o.*
           3s already arrive-PST
           'He has already arrived.'

       c. *tud papɛl prõt jə   fez,   də ĩdian paspɔrt tud jə   dew.*
           all paper ready already make.PST of indian passport all already give.PST
           '[I] have already prepared all papers, [I] have already submitted the Indian passport and all that.'

       d. *ɔr   nə   fik-o   ain.*
           hour NEG$_{cl}$ become-INF still
           'It's not time yet.'

Frequency adverbs also tend towards a preverbal slot, but they can easily occur in a different position. Notice the placement of the adverb *otrez* 'again' in (229):

(229) a. *diŋer   a   gaŋ-a,   faz-e   kaz,   dəpəy kaz-a,   dəpəy otrez*
money IRR.NPST earn-INF make-INF house after marry-INF after again
a   v-i   Diw.
IRR.NPST come-INF Diu
'[I] will earn money, make a house, then get married, then come to Diu again.'

b. *oj   otrez el də afternoon tə   diz-e   a   mĩ   ki*
today again 3s of afternoon IPFV.NPST say-INF DAT 1s.OBL CMP
bam   faz-e   kwɔlkɛ chaknũ
HORT make-INF any   snack
'Today again they told me: "Let's make some snacks".'

In (229a), the adverb occurs immediately to the left of the VP, whereas in (229b) they are mediated not only by an argument but also a temporal PP.

In the corpus, sentence adverbs always occur clause-initially, as demonstrated in (230):

(230) a. *pədser yo vay   Una amiŋã.*
perhaps 1s go.NPST Una tomorrow
'Maybe I will go to Una tomorrow.'

b. *parɛs   use t-iŋ   nə Una õt.*
apparently 2s EXS-PST LOC U.   yesterday
'It seems you were in Una yesterday.'

7.7   Negation

DIP distinguishes between clausal negation (NEG$_{cl}$), propositional (NEG$_p$) negation, and constituent (NEG$_{ct}$) negation.

Clausal negation is operated by the negative particle *nã* (sometimes reduced to *nə*), immediately preposed to the main verb and auxiliaries. No further syntactic means are necessary for the purpose, and therefore DIP negation can be described as *symmetric* (following the typology described in Miestamo 2005). Compare the affirmative sentence and its negative counterpart in (231):

(231) a. *yo kɛ   fal-a   mem, pɔd   fal-a.*
1s want.NPST speak-INF EMPH can.NPST speak-INF
'I really want to speak, [I] can speak.'

b. *yo nã   kɛ   fal-a   mem, nã   pɔd   fal-a.*
1s NEG$_{cl}$ want.NPST speak-INF EMPH NEG$_{cl}$ can.NPST speak-INF
'I really don't want to speak, [I] cannot speak.'

209

The sentences in (232) include various instances of clausal negation:

(232) a. *aki elz nã faz bẽ kim-a, elz nã lav kar*
here 3p NEG$_{cl}$ make.NPST well burn.INF 3p NEG$_{cl}$ wash.NPST meat
dret.
properly
'They don't roast it well here, they don't wash the meat properly.'

b. *yo nã kɛr kõt-a, dig a el.*
1s NEG$_{cl}$ want.NPST tell-INF say.NPST DAT 3s
'I don't want to tell, ask him.'

c. *də Go nã mãd-o Purtəgal, ot lad mem dix-o.*
of Goa NEG$_{cl}$ send-PST Portugal other side EMPH leave-PST
'[They] didn't send [it] from Goa to Portugal, [they] left it some other place.'

The particle for propositional negation, used for instance in reply to polar and tag questions, is *nãw*; notice that, even though the particle for clausal negation is a reduction of this, the unreduced form is hardly ever used in clausal negation (the reason for which it is classified as an interjection in 6.9.2) and the reduced form is almost entirely disallowed in propositional negation. An example of both can be found in the answer of the following short dialogue:

(233) - *are Ashley, use tə fik-a med də cockroach?*
INTJ A. 2s IPFV.NPST become-INF fear of cockroach
'Hey Ashley, are you scared of cockroaches?'

- *nãw, yo nə tə fik-a.*
NEG$_p$ 1s NEG$_{cl}$ IPFV.NPST become-INF
'No, I´m not.'

The crucial distributional variable is whether or not the negative particle is followed by an intonational break. In contrastive assertions such as the ones in (234a), for instance, context determines the omission of part of the clause, and in these cases the negative particle occurs sentence-finally. Such is also the case in (234b), because the predicate has already been established in the previous clause. In such instances, the particle takes its heavy form *nãw*:

(234) a. *a cham-a pə mĩ nɛt, mĩ subriŋ, ot*
IRR.NPST call-INF DAT 1s.POSS grandchild 1s.POSS nephew other
jẽt də for nãw.
people of outside NEG$_{cl}$

'I will invite my grandchilfren, my nephews, other outside people not.'

b. nã   fal   asĩ, nãw, asĩ kufɔr yo fal   ku   syor
   NEG$_{cl}$ speak.NPST thus NEG$_p$ thus like   1s speak.NPST COM mister
   nãw.
   NEG$_{cl}$
   '[We] don't speak like this, no. Not the way I'm speaking to you, no.'

Constituent negation involves the negative indefinite pro-forms *nad* 'nothing', *nĩge* 'nobody' and *nuk* 'never' (see 6.6.5), as well as the NP negator *niŋũ* 'no [X]'. All these elements require subsequent clausal negation, as DIP displays consistent negative concord. Some examples of constituent negation involving the NP negator can be found under (235). Notice that, in (235b), the head of the NP is omitted, so the NP negator appears on its own:

(235) a.   *yo nã   te   niŋũ   amig.*
            1s NEG$_{cl}$ have.NPST NEG$_{ct}$ friend
            'I have no friends.'

       b.   [Which banana do you want?]

            niŋũ   nã   kɛr.
            NEG$_{ct}$ NEG$_{cl}$ want.NPST

            'I don't want any.'

Negative indefinite pro-forms are equivalent to negated NPs, with the exception that reference is indefinite and the negation is universal. An example of each is provided in (236):

(236) a.   *nĩge   nã   apĩŋ-o   pex.*
            nobody NEG$_{cl}$ catch-PST fish
            'Nobody caught fish.'

       b.   *el nuk   nə   foy   sinem.*
            3s never NEG$_{cl}$ go.PST cinema
            'He never went to the movies.'

       c.   *tɛ   agɔr nã   vey   nad.*
            until now NEG$_{cl}$ come.PST nothing
            'Up until now, nothing has come.'

Negative concord is obligatory in DIP to the extent that a negative pro-form will command NEG_{cl} even in the event of the main verb being left unspecified, as shown in the dialogue in (237):

(237)   -   *uki   fik-o?*
            what become-PST
            'What happened?'

        -   *nad   nãw.*
            nothing NEG_{cl}
            'Nothing.'

7.8   Pragmatic status marking

This section explores the formal strategies by which the pragmatic status of particular constituents is conveyed. Some of these involve the alteration of the canonical word order (described in section 7.8.1) - which is, in fact, seldom fully actualised in actual speech -, others imply the null (section 7.8.2) or doubled (section 7.8.3) expression of constitutents. Others still - such as prosodic modulations (section 7.8.4) or the use of pragmatic status markers (section 7.8.5) do not involve any such changes.

The compartmentalisation of this discussion into subsections does not mean to imply that the various factors do not interact. Ellipsis of superfluous or secondary constituents, for instance, often brings about contingent changes to the prototypical order of the constituents which speakers do realise. Similarly, the operation of pragmatically-induced word order changes does not prevent the additional use of focus prosody or pragmatic status markers.

7.8.1   Word order

One of the crucial factors to understand pragmatically-motivated word order changes is the recognition of a focus position, which is preverbal in DIP. The associated ordering principle is paraphrased as follows:

- Ordering principle 1: Express particularly prominent constituents in preverbal position.

I will begin to illustrate the workings of a preverbal focus position with reference to monotransitive clauses, although the same applies to the U and R arguments of ditransitives as well. When particular focus is intended on the P argument of a monotransitive clause, this argument will display a tendency towards a preverbal position (i.e., prior to the verb and its modifiers), a sort of fronting strategy. This tendency, however, often clashes with a preference for not juxtaposing the two arguments of a transitive clause before the

verb, according to another motivation which can be paraphrased as follows:

- Ordering principle 2: Do not express more than one argument in preverbal position.

One common resolution to this clash, in DIP, is the repetition of the predicate. For a description of the phenomenon, see section 7.8.3.

Examples of P-A-V are few and far between. The sentences in (238) effectively place P in front of A, while APV is not represented at all in the corpus except in copular constructions:

(238) a. battery *aŋe*   *rub-o*   *foy.*
           battery someone steal-PST go.PST
           'Someone stole the battery and left.'

       b. *tud yo sab   faz-e.*
           all 1s know.NPST make-INF
           'I can do everything.'

       c. *es   chɛf Rujɛr de-w.*
           DEM$_p$ chief R.   give-PST
           'Chief Rogério gave [me] this.'

       d. *pə   Bablu kẽ   vay   cham-a?*
           DAT B.   who go.NPST call-INF
           'Who will go call Bablu?'

The operation of a preverbal focus position can also be observed in ditransitive constructions. This is shown in (239), in which the R argument is assigned contrastive or exhaustive focus:

(239) *a   el yo de-w,   mayz kẽ?*
       DAT 3s 1s give-PST more who
       'It was to him that I gave [it], who else?'

Not only arguments, but also adjuncts and subordinate clauses, may be placed in preverbal position for focus enhancement. Given that the position of adjuncts is in general more flexible than that of arguments, the pull of the focus position is less evident, but equally operative. In the following examples, notice how a source PP (240a) and a purpose adverbial clause (240b) appear in preverbal position on account of their centrality to the message conveyed:

(240) a. *də nĩge yo nə   tə   fik-a   med.*
of nobody 1   NEG$_{cl}$ IPFV.NPST become-INF fear
'I am not scared of anyone.'

c. *karner pə   asĩ faz-e   nã   sɛrv   pərki el*
mutton PURP thus make-INF NEG$_{cl}$ be.adequate.NPST because 3s
ɛ   dur.
COP$_i$.NPST hard
'Mutton is not good to do this because it is hard.'

Mere alteration of the prototypical word order is not the only focus-attributing strategy available to speakers of DIP. There is a construction similar to a cleft - with the crucial distinction that only the relative clause contains a verbal form - which assigns contrastive focus onto the clefted argument. Some examples are given under (241) - the clefted constituents are given in square brackets:

(241) a. *[yo] ki fez   es.*
1s   REL make.PST DEM$_p$
'It was me who did that.'/'I did that myself.'

b. *ikəl   tud koyz ki lɛv-o   nə   museum [father Marian] ki*
DEM$_d$ PL thing REL take-PST LOC museum father M.   REL
lɛv-o   nə?
take-PST REQ
'Wasn't it father Mariano who took those things to the museum?'

The alternative translations of (241a) indicate that the focus placed on the agent by this cleft construction may impart a reflexive reading. In that respect, the construction is functionally equivalent to the use of the emphatic particle *mem* (see section 7.8.5).

Section 7.8.2 explores the interaction between ellipsis, focus and word order, namely in determining the preverbal or postverbal placement of particular constituents in clauses from which other constituents are omitted.

7.8.2   Ellipsis

DIP places few restrictions on the omission of arguments which would be expected from the semantic valency of certain verbs. Let us take the typically ditransitive verb *da* 'to give' as an example. The action described by this verb form typically involves an agent, an undergoer and a recipient (see section 7.5) to which various adjuncts can be appended. In DIP, the manifestation of all these elements is very much optional and takes on board pragmatic variables such as the referential retrievability or unimportance of the constituents; there is a logical distinction between *zero-anaphora* (when the unrealised element is obvious

from the immediately preceding discourse context and/or shared knowledge) and *valency-decreasing omission* (when ellipsis is motivated by the irrelevance of a clausal constituent), and it is clear that in DIP both motivate ellipsis. Content questions (see section 7.6.2.3) easily elicit elliptical formulations not only because they can flag part of the predication as given information, but also because they highlight the importance of a constituent above the others. To return to the example of the verb *da*, consider the sentences in (242), in which different questions result in the focused realisation of an agent (242a), an undergoer (242b), a recipient (242c) and a temporal adjunct (242d):

(242)   a.   [Who gave you that?]

daddy (*de-w*).
daddy give-PST

'Daddy gave [it to me].'

b.   [What did he give you?]

tud (*de-w*).
all   give-PST

'[He] gave [me] everything.'

c.   [Who did you give your bicycle to?]

a   *el* (*de-w*).
DAT 3s give-PST

'[I] gave [it] to him.'

d.   [When did your daddy give you the bicycle?]

nə   *mĩ*   birthday (*de-w*).
LOC 1s.POSS birthday give-PST

'[He] gave [it to me] on my birthday.'

The brackets surrounding the verb forms in (242) indicate that the verb itself is optional in such cases. In (243), the copula in the second clause is omitted as well as the argument. As far as the verbal form of the first sentence is concerned, its intepretation as an intransitive existential or montransitive possessive verb rests on whether one assumes an ellipted possessor (cf. sections 7.1.3.3 and 7.1.3.4):

(243)   *kume nã   te   nə? pɔb mem   nə?*
food   NEG$_{cl}$ *te*.NPST REQ poor EMPH REQ
'There is no food/[I] have no food, you see? [I am] very poor, you see?'

It is also clear that the absence of certain constituents may affect the prototypical place-ment of those realised. This is particularly clear in (242b,c) above, in which the undergoer and the recipient occur preverbally rather than postverbally. The preverbal placement of these arguments is, however, variable and optional, and it seems to be governed to some extent by notions of focus placement (for which, see 7.8.1). Notice, for instance, the fol-lowing examples:

(244)   a.   [What did you do in Goa?]

        kume-w   churis.
        eat-PST sausage.

        '[I] ate sausage.'

    b.   [What did you eat in Goa?]

        churis   kume-w.
        sausage eat-PST

        '[I] ate sausage.'

In (244a), the question places equal emphasis on the undergoer and the action, and as such the elements occur in their basic order despite the absence of an agent argument. In (244b), on the other hand, the action is provided by the question and particular emphasis is therefore placed on the undergoer, which appears in preverbal position in the answer. It must be clarified, however, that pragmatic effects of this kind are tendential rather than deterministic when it comes to the syntactic position of the elements expressed.

The omitted elements are sometimes appended at the end of the clause as afterthoughts, in which case they are preceded by an intonational break. This is shown in (245):

(245)   a.   *es   abri-w,   use?*
        DEM$_p$ open-PST 2s
        'Did you open this?'

    b.   *fik   kalad, use.*
        stay.NPST silent 2s
        'Shut up, you.'

    c.   *vay   vist-i   nə,   use.*
        go.NPST dress-INF REQ 2
        'Go get dressed, you.'

The examples in (245) include both interrogative (245a) and imperative (245b,c) clauses.[31] (246) exemplifies further instances of argument ellipsis within the context of interrogative clauses. Unless overruled by discourse context, the intended subject is usually the interlocutor:

(246) a. *kwɔn kõpr-o?*
           when buy-PST
           'When [did you] buy [it]?'

       b. *mis foy?*
           mass go.PST
           '[Did you] go to mass?'

The relative lack of restrictions on ellipsis poses a theoretical question concerning adnominal interrogatives. These sometimes appear in non-attributive position, so one is faced with the decision whether to assume a headless NP or to classify them alongside clause-level interrogative pro-forms. The issue will not be resolved here, but this use of *kwɔl* 'which' and *kwõt* 'how much/many' is exemplified in (247):

(247) a. *ũ   istɔr... kwɔl a   kõt-a?*
           one story   which IRR.NPST tell-INF
           'A story... which (one) should I tell?'

       b. *use kwõt   prisiz?*
           2   how.much need.NPST
           'How much do you need?'

Adnominal interrogatives and their usual syntactic context are described in section 7.2.4.

7.8.3 Constituent doubling

Constituent doubling is very common in DIP, though strictly optional. Various types of constituents can be doubled, and the functions of this strategy are various. Even though it is often difficult to ascertain beyond doubt the reasons behind each instance of constituent doubling, the motivation seems to be entirely pragmatic. I will begin by describing cases

---

[31] In (245c), the clause boundary - after which the pronominal argument is added - is marked by the requestative *nə*, but this is not enough to justify the extraclausal interpretation of the pronominal argument, because requestatives/tags may either occur clause-finally or immediately after the verb (cf. sections 7.6.2.2 and 7.6.3.1). In this case, apart from an intonational boundary (graphed as a comma), we also have the atypical syntactic placement of the pronominal argument to justify this interpretation.

of verb doubling (not to be mistaken with verb reduplication, described in 6.2), which are the most common.

Whenever both A and P are realised, the outcome is overwhelmingly A-V-P, whereas the realisation of both A and U/R results in A-V-U/R. However, if P, U or R receive pragmatic focus, they will ideally occupy a preverbal position (see 7.8.1). The corpus of free-flowing speech contains several clauses in which this clash is resolved by repeating V, giving rise to such structures as A-V$_1$-P-V$_1$, A-V$_1$-R-V$_1$ or A-V$_1$-U-V$_1$. It is clear that the elements preceding the doubled verb are particularly prominent, as they typically constitute the intonational peak of the clause (see 7.8.4). Example (248a) was uttered by a young boy, whereas (248b), spoken by a teenage girl, is a declarative sentence turned into a polar question by means of the tag (requestative) *nə* (see section 7.6.2):

(248) a. *sĩ, ĩglix   fal,   may faz   mistakes faz.*
           yes English speak.NPST but make.NPST mistakes make.NPST
           'Yes, he speaks English but he makes mistakes.'

       b. *ãt   kastɛl ɛ-r   prizãw ɛ-r   nə, kaʃtɛl?*
           before fort   COP$_i$-PST prison COP$_i$-PST REQ fort
           'Earlier the fort was a *prison*, wasn't it?'

Whereas the avoidance of argument juxtaposition may be a strong motivation behind verb doubling, the truth is it also occurs when some constituents are elyptically omitted. Notice the sentence in (249):

(249)   [Who did you give it to?]
           de-w   a   el de-w.
           give-PST DAT 3s give-PST
           'I gave it to him.'

In (249), the non-realisation of the A argument leaves the preverbal position unoccupied, so in theory there is no reason why the R should not appear clause-initially as it does elsewhere in similar cases. It is clear, in addition, that the question attributes higher prominence to R than to V (which is given), and this - in accordance with Ordering principle 1 above - justifies the tendency to have R in preverbal position. In such cases, one probably has to assume yet another competing tendency entering the equation, paraphrased as follows:

- Ordering principle 3: Express arguments in their prototypical position.

Apart from predicates, various other elements can be doubled. The double occurrence of the restrictive particle *sɔ* 'only' in (250) may seem logically redundant, but its doubling effectively strengthens the intended exhaustive focus (for the typical use of *sɔ*, see 7.8.5).

This is further enhanced by the subsequent clause, which asserts that another logical alternative, viz. that the speaker went to Daman in addition to Mumbai, is false:

(250) *yo foy   sɔ   Bomə̃ĩ   sɔ,   Damã̃w nã   foy.*
       1s go.PST only Bombay only Daman NEG_{cl} go.PST
       'I went to Bombay only, I didn't go to Daman.'

Arguments may also be repeated, in which case the function is simply one of additional emphasis or contrastive focus, as exemplified in (251a). This is particularly common as a means to add prominence to a constituent in polar questions (251b) and tag questions (251c):

(251)   a.   *a   mĩ̃   el de-w,   a   mĩ̃.*
              DAT 1s.OBL 3s give-PST DAT 1s.OBL
              'He gave [it] to me.'

         b.   *use aŋũ vez   jə   foy   Damã̃w, use?*
              2   any occasion already go.PST Daman   2
              'Have you ever been to Daman, you?'

         c.   *use ain nã   t-iŋ   aki nə,   use?*
              2   yet NEG_{cl} EXS-PST here REQ 2
              'You weren't here yet, were you?'

The sentences in (251) contain a strong intonational break before the repeated element, so that we may consider these sequences extra-clausal - see also section 7.8.2. In that respect, then, this type of argument doubling differs from verb doubling as described above.

7.8.4   Prosody

The contribution of prosody to pragmatic status marking is rather straightforward. The prosodic contour of a clause may highlight focused constitutents, as these often attract the intonational peak of the clause. The assignment of prosodic prominence often goes hand in hand with other mechanisms of pragmatic marking (word order alteration, constituent doubling, ellipsis), but one cannot say that prosodic focus marking is the cause of syntactic focus marking, or vice-versa. In general, it appears that prosody is not as determinant a factor for pragmatic status marking in DIP as it is in other languages, given that it does not compete with or disallow other focussing strategies.

For further information on DIP intonation, see section 5.4.

7.8.5 Pragmatic status markers

The particle *mem* is extremely common in DIP and occurs in various positions, with scope over various clausal and phrasal elements. This particle has a generic emphatic function, with the peculiarity of being allowed to modify any element, both clausal and phrasal - which means *mem* can be internal to a phrase. This particle is usually postposed to the element(s) under its scope. (252) exemplifies several occurrences of *mem* modifying various constituents of NPs and AdjPs; notice how the particle attaches to a noun in object position (252a), an attributive possessor (252b), a predicative pronominal possessor (252c), a predicative adjective (252d,e)[32] and an intensifier (252f):

(252) a. *yo sa b-iŋ   ki   el tə   fal-a   Liza mem.*
           1s know-PST CMP 3s IPFV.NPST say-INF L.   EMPH
           'I knew that he would say "Liza" only.'

     b. *kurəsãw də makak dẽt   d-el mem   korp.*
           heart   of monkey inside of-3s EMPH body
           'The heart of the monkey [is] inside its own body.'

     c. *nãw, es   ε   də mĩ   mem   ε.*
           NEG$_p$ DEM$_p$ COP$_i$.NPST of 1s.OBL EMPH COP$_i$.NPST
           'Now, this is my own.'

     d. *vεy mem   nə?*
           old EMPH REQ
           '[He is] really old, isn't he?'

     e. *es   ε   mej~mej   mem, sĩ, mej mem   ε.*
           DEM$_p$ COP$_i$.NPST same~same EMPH yes same EMPH COP$_i$.NPST
           'This is exactly the same, yes, it's the very same.'

     f. *bẽy mem   pɔb ε.*
           very EMPH poor COP$_i$.NPST
           '[I] am really very poor.'

AdvPs and adverbial clauses also admit this element. In the examples below, *mem* emphasises a manner adverb (253a) and a locative adverb in an appositional sequence (253b):

---

[32] The combination of *mej* 'same' and *mem* 'EMPH' in (252e) is very interesting because both words derive from the same Portuguese etymon *mesmo* - which, however, is used in Portuguese as both an adjective 'same', an intensifier 'really' and a reflexive particle '-self'. It is clear that, in the history of DIP, functional segmentation resulted in formal segmentation, so that nowadays *mej* and *mem* are separate forms with separate functions.

(253) a. *asĩ   mem   nə   Purtəgal jə   vir-a   nom?*
           thus EMPH LOC Portugal already turn-INF name
           'In Portugal, do [people] change [Lit. 'turn'] the names just like this?'

     b. *aki   mem   pɛrt, aki.*
         here EMPH near here
         'Right nearby [Lit. 'close to here'].'

Verbs or entire predicates can also fall under the scope of emphatic *mem*, as shown in (254):

(254) a. *yo nã   kɛr   fal-a   mem.*
           1s NEG$_{cl}$ want.NPST speak-INF EMPH
           'I really don't want to speak.'

     b. *nã   de-w   mem   rɛpos.*
         NEG$_{cl}$ give-PST EMPH answer
         '[They] didn't answer at all.'

     c. *nad   nã   tə   fik-a   mem   asa-d   dret.*
         nothing NEG$_{cl}$ IPFV.NPST become-INF EMPH roast-PTCP properly
         'Nothing ever gets roasted properly.'

To the extent that identifiability and referentiality are context-dependent, the markers of definiteness - which are equivalent to the demonstrative series (in the case of definite marking) and the numeral 'one' (for indefinite reference) - are also considered pragmatic markers. These are described in detail in section 7.2.6.

Likewise, the restrictive particle *sɔ* 'only' is instrumental in establishing the boundaries of the intended reference and often also to highlight either exhaustive focus or the paucity of a given referent. This particle occurs both before and after the element it modifies, which can be a headless NP (255a), a nominal head (255b) or a verb (255c):

(255) a. *pok   sɔ   sab.*
           little only know.NPST
           '[He] knows only a little bit.'

     b. *nãw, asĩ nã   fik,   buliŋ   sɔ   fik.*
         NEG$_p$ thus NEG$_{cl}$ become.NPST croquette only become.NPST
         'No, [mutton] won't be good like this, only as croquettes.'

c. *sɔ mw-er   a   el, dəpəy põy   sum də limãw.*
only grind-INF DAT 3s then put.NPST juice of lime
'[You] only [have to] grind it, and then add a little lime juice.'

The logical counterpart of the restrictive particle *sɔ* is the additive particle *time* 'also', which is more consistently postposed to the element(s) under its scope. Therefore, in (256a) *time* will normally be understood as adding a participant to the action, whereas in (256b) it adds an action to a participant (in the latter case, the speaker had previously declared he had spent the previous evening afternoon riding his bike):

(256) a. *nɔs time fik   nə   pulis* quarters.
1p   also stay.NPST LOC police quarters
'We also live at the police quarters.'

b. *yo tə   ve   tv time.*
1s IPFV.NPST see.INF tv also
'I watched television as well.'

7.9 Coordination

A construction involving two elements of the same type (e.g. two NPs or two clauses) is considered coordinated if they have a roughly equivalent status, i.e. if there is no perceived asymmetry in their salience or importance. Following the terminology proposed in Haspelmath (2004), I will use *coordinand* to refer to any of the coordinated elements and *coordinator* to refer to the grammatical element which marks the coordination. In DIP, coordination is mostly *monosyndetic* - i.e., it overtly marks the coordination with one coordinator only - and the coordinator, which is always a free word, occurs in between the coordinands - it is therefore a *medial connective* in the terminology of Stassen (2005b). In the rare eventuality that the coordinands are separated by external syntactic material (i.e., in the extraposition of one of the coordinands), the coordinator always occurs just before the second coordinand - see (263). DIP coordination can therefore be represented with the schema in (257), in which C_1 and C_2 represent two coordinands, *co* stands for the coordinator and X stands for external material (the parentheses indicate optionality):

(257) $[C_1]$ (X) [co C_2]

Apart from their semantic differences, the various coordinators in DIP reveal some degree of category-sensitivity. For instance, the conjunctive coordinator *ku* is allowed as a linker of NPs only, whereas the conjunctive coordinator *i* may conjoin NPs, VPs, clauses and more. In this description, I will follow the traditional semantic macro-divisions of coordination, viz. conjunctive, disjunctive and adversative.

7.9.1   Conjunctive coordination

Conjunctive coordination typically has an additive function in that all coordinand proposi-
tions are expected to have the same truth value. DIP makes use of different conjunctions,
with different distributional characteristics. As mentioned above, the conjunction *ku* 'and'
- which is equivalent to the comitative and instrumental marker *ku/ko* 'with'[33] (see section
7.4.7) - is strictly restricted to NP coordination. Given the macrofunctionality of *ku*, it is
often unclear whether the intended reading is conjunctive, comitative or instrumental. Let
us first look at the examples in (258):

(258)   a.   *oj   ε   nɔs   exam *gujərati   ku *ĩglix.*
             today   COP$_i$.NPST   1p.POSS   exam Gujarati   *ku* English
             'Today is our Gujarati and English exam.'

        b.   crocodile *ku makak   Steven kõt-o   istɔr.*
             crocodile *ku* monkey   S.   tell-PST story
             'Steven told the story of the crocodile and the monkey.'

        c.   *kãw ku* lion *kuməs-o   jug-a.*
             dog *ku* lion begin-PST play-INF
             'The dog and the lion started playing.'/'The dog started playing with the lion.'

        d.   *fil   də mĩ   ku nɛt,   ɛl fik   ali   nə   OIDC.*
             child of 1s.OBL *ku* grandchild 3f stay.NPST there LOC OIDC
             'My daughter lives at the OIDC with my grandchildren.'

In (258a), coordination is the most obvious reading, given that the coordinands are
inanimate and abstract. In (258b), there is also no reason not to interpret the function of *ku*
as coordination, in particular because in the story referred to the monkey is clearly not the
least prominent character (much on the contrary). The situation is more complex in (258c);
all competing readings are conceivable, and there are no formal means of distinguishing
one from the other. Finally, the comitative reading would be dominant in (258d) if one
could be certain that the resumptive pronoun *ɛl* is singular, but the corpus does show
that the morphological marker of plurality on 3rd person personal pronouns is optional
and not always activated for plural reference. Considering that the working definition of
coordination requires the coordinands to have similar prominence and status, the issue of
whether or not the coordinands in these sentences are absolutely symetrical is not trivial
(for a related discussion, see Gil 2004, Haspelmath 2004, Stassen 2005a). It is nonetheless
impossible to resolve it here. The point to be retained is that the semantic scope of *ku*
encompasses both conjunctive and comitative, among others.

---

[33]The isomorphism of NP-coordinators, comitative markers and instrumental markers is very common,
cross-linguistically (see e.g. Payne 1997, Haspelmath 2004, Stassen 2005b). Stassen (2005a) also makes
the point that, in languages displaying this isomorphism, the comitative-like marker is often not the only
NP-coordinator allowed.

The element *may*, which doubles as a quantifier/intensifier meaning 'more', is semantically very close to *ku*, to the extent that the same difficulty in distinguishing between a coordinating or comitative reading applies. Consider sentence (259):

(259)   *yo may doy rapas tə*   *trabəy-a.*
       1s *may* two boy   IPFV.NPST work-INF

       'Me and two boys work [here].'/'I work with two boys.'

May is much less common in such contexts than *ku*, and either of these less so than the conjunction *i* 'and'. The conjunction *i* easily replaces NP-coordinating *ku*. In order to illustrate this, I repeat a small dialogue in (260), which records their equivalence. *C* indicates a collaborator, who uses both *i* and *ku* in precisely the same context, and *R* indicates the researcher. This excerpt was previously transcribed in 2.3:

(260)   C   *nə*   *Go* *yo te*   *bastãt* cousin *i* auntie.
         LOC Goa 1s have.NPST many cousin *i* auntie

         'I have many cousins and aunties in Goa.'

    R   [te bastãt . . . ?]

    C   *cousin ku auntie.*
         cousin *ku* auntie

         'Cousins and aunties.'

Compare also (261) with (258a) above:

(261)   *ĩglix*   *i*   *gujərati* exam *kab-o.*
       English and Gujarati exam finish-PST

       'The Gujarati and English exam is over.'

The examples in (262) include further instances of NP-coordinating *i*:[34]

(262) a. *el sab   mariti   i   ĩglix.*
       3s know.NPST Marathi and English
       'He knows Marathi and English.'

    b. *Conchita i   Balfina ɛ   mɛdroz.*
       C.   and B.   COP$_i$.NPST fearful
       'Conchita and Balfina get scared very easily.'

    c. *Liza, Ashley, Fabian i   Conchita tə   faz-e   racing.*
       L.   A.   F.   and C.   IPFV.NPST make-INF racing
       'Liza, Ashley, Fabian and Conchita are racing.'

Sometimes, what looks like coordinated NPs may occur discontinuously, as demonstrated in (263). In these cases also, the coordinator occurs next to the final coordinand:

(263) *ũ   makak t-iŋ   i   ũ   crocodile.*
       one monkey EXS-PST and one crocodile
       '[Once upon a time], there was a monkey and a crocodile.'

In addition to NPs, *i* also links different types of coordinands, such as clauses and VPs. I will begin by discussing VP coordination, because it raises some questions concerning the status of what could be interpreted as serial verb constructions.

In DIP, VPs with the same subject may be coordinated with an overt conjunctive coordinator *i* 'and', as in (264a) or without a coordinator as in (264b). In truth, these are cases of simple clausal coordination, with the peculiarity that the shared constituents of the various clauses are not repeated - or they may be left unexpressed altogether, in case it is sufficiently clear from (discursive) context. In order to distinguish such cases of conjunctive coordination from those in which the various clauses are entirely independent

---

[34]Compare (262a) with the following sentence:

(1) *yo sab   gujərati, yo sab   ĩglix   i   yo sab   purtəgez.*
   1s know.NPST Gujarati 1s know.NPST English and 1s know.NPST Portuguese
   'I know Gujarati, I know English and I know Portuguese.'

Their logical similarity is evident, and yet one speaks of NP-coordination in (262a) and of clausal coordination in the sentence above. The difference rests solely on the amount of redundancy allowed by the speaker in one or the other, and it has been recognised before (e.g. Haspelmath 2004) that one may prefer to interpret constituent coordination as clausal coordination with elliptical omission of the shared constituents. In that sense, for instance, (262c) would be a variant of *Liza tə faze* racing *i Ashley tə faze* racing *i Fabian tə faze* racing *i Conchita tə faze* racing. This discussion is rather trivial for DIP, and therefore I will make no strong claims concerning the theoretical controversy. Whenever I mention *NP-coordination* or *VP-coordination*, for instance, the labels are purely formal.

from each other, I will term the phenomenon *predicate enumeration*.[35]

(264)   a.   *use vay   i   trag.*
                  2s   go.NPST and bring.NPST
                  'Go and bring [it].'

       b.   *vay   tɔm.*
                  go.NPST take.NPST
                  'Go and get [it].'

In these examples, both coordinated verbs are Imperative.[36] Tense and aspect information usually does not travel through clausal frontiers in predicate enumerations - but see (273) below; as a result, all verbs have to be separately marked for tense and aspect categories even though they often coincide, as is the case in (265) below.

(265)   *yo a   faz-e   prõt (i) a   dix-a.*
       1s IRR.NPST make-INF ready and IRR.NPST leave-INF
       'I will prepare it and (I will) leave it.'

Predicate enumerations may involve several elements. This is demonstrated in (266), which is also a clear example that enumerated predicates are often ordered chronologically.

(266)   *foy   i   pasy-o   pray, kõpr-o grãw i   vey   kaz. dəpəy*
       go.PST and take.walk-PST beach buy-PST gram and come.INF house then
       vey   kaz,   rɛz-o   ters i   kume-w i   durmi-w.
       come.PST house pray-PST beads and eat-PST and sleep-PST
       '[We] went and took a walk around the beach, bought some gram and came home. Then [we] came home, counted our beads, ate and went to sleep.'

Notice that, in the previous sentence, *i* occurs sporadically within the enumeration. There is however a strong tendency to mark only the last coordinand, as demonstrated in (267). In this sentence, the first *i* is a sentence coordinator, or discourse marker; the final clause *lion falo ki muyt ubrigad* is clearly not part of the predicate enumeration any more, given that its subject is not the same as that of the previous clauses:

---

[35] The term *clause chaining* is deliberately avoided because it has been applied in the literature to refer to very concrete cases of languages in which the various components of a clause chain are somewhat grammatically integrated, in that only one of them (usually the last one in the sequence) receives full TAM marking. Such systems blur the coordination/subordination divide. The situation is rather different in DIP; in the constructions I call *predicate enumeration*, the verbs in the different clauses are separately marked for tense and aspect.

[36] For an account of the imperative reading of these sentences, see section 7.6.3.

(267) *i   rat   foy,   murde-w murde-w pikǝliŋ, sult-o   i   dǝpǝy* lion
and mouse go.PST bite-PST bite-PST little   release-PST and then lion
fal-o   ki   muyt ubrigad.
say-PST CMP much thanks
'Then the mouse went, [he] bit for a while, [he] released [the lion] and then the lion said: "Thank you very much".'

The fact that the coordinator *i* is strictly optional in predicate enumeration makes it unclear whether to interpret instances of predicate enumeration without an overt connector - such as (264b) above and (268a,b) - as serial verbs.

(268) a. *i   dǝpǝy ũ   lion vey   subi-w   ali   isim.*
and then   one lion come.PST climb-PST there on.top
'And then a lion came and climbed up there.'

    b. battery *aŋe   rub-o   foy.*
battery someone steal-PST go.PST
'Someone stole the battery and left.'

Consider, for this discussion, the following definition of serial verb constructions:

> A serial verb construction (SVC) is a sequence of verbs which act together as a single predicate, without any overt marker of coordination, subordination, or syntactic dependency of any other sort. Serial verb constructions describe what is conceptualized as a single event. They are monoclausal; their intonational properties are the same as those of a monoverbal clause, and they have just one tense, aspect, and polarity value. SVCs may also share core and other arguments. Each component of an SVC must be able to occur on its own. Within an SVC, the individual verbs may have same, or different, transitivity values. (Aikhenvald and Dixon 2006:1)

A DIP sentence such as (264b) or (268) seem to fulfill all of these requirements. However, in typical descriptions of serial verbs systems, a given construction should be classified as a serial verb construction only if no alternation with predicate coordination in exactly the same syntactic contexts ever occurs. It is therefore merely a theoretical matter whether or not to take the optional occurrence of *i* in these contexts as proof that serial verb constructions are absent from the syntactic array of DIP.

In addition to the predicate enumerations discussed above, there are also instances of clause coordination in which the coordinand clauses share the subject. Some examples are given in (269):

(269) a. *el fik-o   ĩtrujad i   el foy   kas.*
3s become-PST fooled and 3s go.PST house
'He was fooled and he went home.'

b. *dez di   el kume-w dəpəy el foy   i   el dis   [. . .].*
   ten day 3s eat-PST then 3s go.PST and 3s say.PST
   'He ate for ten days, then he went and he said [. . .].'

The coordinand clauses may also have no shared elements at all, as demonstrated in (270):

(270) *makak   vey   i   crocodile dis   yo kɛr   kum-e d-use kurəsãw.*
       monkey come.PST and crocodile say.PST 1s want.NPST eat-INF of-2   heart
       'The monkey came and the crocodile said: "I want to eat your heart."'

It is interesting to notice that, once more, the coordinand clauses are ordered chronologically, simultaneously indicating the completion of one action and the the inception of a subsequent one - this is the *sequential* function associated wih conjunctive coordination in many languages (Malchukov 2004). This use of *i* is therefore similar to that of other discourse structuring elements such as *dəpəy*[37] 'then' - which often combines with *i*, such as in (268a) - as exemplified in (271):

(271) *dəpəy chig-o   tɛ   arv nə, rabbit kãs-o,   dəpəy el drumi-w   ali.*
       then arrive-PST until tree REQ rabbit tire-PST then 3s sleep-PST there
       'Then [he] arrived at the tree, you see, the rabbit got tired, then he fell asleep there.'

7.9.2   Disjunctive coordination

In disjunctive coordination, the relationship between the coordinand propositions is such that, for the coordination to apply, only one of the coordinands can hold true. The DIP disjunctive coordinator is *o* 'or', which can link various types of coordinands. Modifier (AdjP) disjunction is exemplified in (272):

(272) *use kɛr   brãk o pret kamiz?*
       2   want.NPST white or black shirt
       'Do you want a white shirt or a black one?'

---

[37] The corpus also contains instances of the word *sənãw* being used as a comparable discourse marker:

(1) *sənãw el dis   ki   on   te   tortoise?*
     sənãw 3s say.PST CMP where EXS.NPST tortoise
     'Then he said: "Where is the tortoise?"'

This use of *sənãw*, initially a subordinator meaning 'if not', is reported in Clements (forthcoming) as an incipient trend in Daman IP, and in fact it is clear that the Diuese children who produce it are those who descend from Damanese and have a strong family connection with Daman - see 2.2.2 for a discussion of observed 'family effects'.

The sentence in (273), on the other hand, exemplifies clausal disjunction:

(273) *nɔs doy a   faz-e   race i   v-er   use si a   sa-i*
       1p two IRR.NPST make-INF race and see-INF 2   if IRR.NPST leave-INF
       pimer o yo si a   sa-i   pimer.
       first   or 1s if IRR.NPST leave-INF first
       'The two of us will race and we'll se whether you leave first or I leave first.'

Notice that the the coordinated clauses (*use si a sai pimer* and *yo si a sai pimer*) both play the same role in the sentence, as they both constitute complements to the verb *ver* 'to see'.

The simple juxtaposition of different NPs, such as *doy [an] trey an* in (274), often has a disjunctive reading as well.

(274) *yo ad   v-i   volta-d   dəpəy də doy trey   an.*
       1s IRR.NPST come-INF return-PTCP after   of two three year
       'I'll come back in two or three years.'

Disjunctive coordination is much less common in DIP than conjunctive coordination.

7.9.3   Adversative coordination

The adversative coordinator in DIP is *may*[38] 'but'. Its distribution is more constrained than that of the conjunctive and disjunctive coordinators described above, as *may* is restricted to linking full clauses. There are certain semantic nuances to the use of this coordinator, which will be demonstrated below.

The sentences in (275) are labelled *concessive* because the coordinated clauses are expected to be logically incompatible.

(275) a. *tud jə   fes   yo may Go, Go n-əd   da   rɛpɔs.*
           all already do.PST 1s but Goa Goa NEG$_{cl}$-IRR.NPST give.INF answer
           'I have already done everything, but Goa won't give an answer.'

       b. *kriãs dig   es kumid ɛ bunit may nɔs dig*
           child say.NPST –   but 1p say.NPST
           es kumid ɛ muyt gustoz.
           –
           'Children say "*es kumid ɛ bunit*" but we say "*es kumid ɛ muyt gustoz*"'.

---

[38]Not to be mistaken for the quantifier/intensifier *may* 'more'. These have widely different functions and derive from different Portuguese etyma: coordinating *may* derives from Ptg. *mas* 'but', whereas the quantifier/intensifier stems from Ptg. *mais* 'more'.

In (275a), the fact that the speaker has done all that was required by bureaucracy is expeted to result in an answer from Goa, but the logical conclusion has not obtained. In (275b), the pressuposition would be that children and adults use the same expressions, but the speaker clarifies that not be the case.

In a sentence such as (276), *may* has an *argumentative* function - termed *contradicting evaluation* in Malchukov (2004) -, in the sense that there is a contrast between the inferences of the coordinated clauses:

(276) *elz kir-iŋ   faz-e   kaz   may d-elz jŭt   nã   t-iŋ   diŋer.*
       3p want-PST make-INF house but of-3p together NEG$_{cl}$ EXS-PST money
       'They wanted to build a house but they didn't have any money.'

The first clause in (276), viz. that X wanted to build a house, would normally invite the inference that they did; on the other hand, the fact that they did not have the necessary money suggests the contrary inference that they did not build a house. This contrast is here expressed with the adversative *may*.

7.10   Subordination

Unlike coordination, the clauses involved in subordination do not all have the same status, to the extent that subordinate clauses can be said to be dependent on (parts of) a main clause. The functions of subordinate clauses vary. They may constitute an argument, in which case we are dealing with *complement clauses*, or have an adverbial function, in which case they are known as *adverbial clauses*. These two types of clauses will be described separately in this section. Finally, I will discuss the syntax of relativisation in DIP.

Stassen (1985) - see also Cristofaro (2005a) - notices that languages treat their subordinate clauses differently with respect to the degree of morphosyntactic independence they are allowed. In some cases, the verbs in subordinate clauses may admit the full range of TAM, i.e. they are formally undistinguishable from verbs used in simple declaratives; these verb forms (and by extension the clause) are called *balanced*. On the other hand, if the verb form in a subordinate clause behaves in some way differently from similar verbs in main declaratives, they are considered *deranked*. In DIP, as described below, we find both balanced subordinate clauses (those which admit TAM operators and inflection, i.e. finite clauses) and deranked subordinate clauses (in which the verb occurs in its infinitival form without any TAM operators).

In DIP, subordinators (be them relativisers, adverbial subordinators or otherwise) may occur in one of two syntactic positions, viz. before the subordinate clause or immediately preceding its VP. Consider the sentences in (277), which exemplify the competing positions for the conditional subordinator *si* 'if' - in this section, subordinate clauses are bracketed off in example sentences, to ease their interpretation:

(277) a. *ũ   banglo   jə   faz   [si yo apiŋ-a   dĩer].*
           one bungallow already make.NPST if   1s find-INF money

'If I find the money, I will build a bungallow.'

b. [*aŋe si věy*]   *aki pɔd*   *fik-a*.
    someone if come.NPST here can.NPST stay-INF
    'If someone comes, [they] can stay here.

Some subordinators (especially complementisers) are less prone to this type of variation, preferring the clause-initial position. Although the exact constraints on variation in this case still needs to be worked out in more detail, one preliminary observation is that subordinators which substitute for a clausal constituent (such as relativisers, for example) are particularly free to occur in either position.

Syntactically, subordinate clauses can be embedded at the clausal level or the phrasal level. Formally, there are also various types, including:

- finite clauses with an overt subordinator;
- finite clauses without an overt subordinator (rare);
- nonfinite clauses with an overt subordinator;
- nonfinite clauses without an overt subordinator.

I will now describe the distribution of these structures across the various subtypes of subordinate clauses.

7.10.1 Complement Clauses

Complement clauses are defined as those which function as the complement of a verb; formally, they may constitute full clauses or, minimally, consists of a single verb form. Indirect-object complement clauses are entirely unrepresented in the DIP corpus, and subject complement clauses only rarely so. One example of a complement clause with a subject function is given in (278):

(278)   [*as-a*   *dret*]   *lɛv*   *pok tẽp nə*.
      roast-INF properly take.NPST little time REQ
      'Roasting properly takes some time, you see.'

The vast majority of complement clauses therefore occupy object slots. The most basic type of complement clause is that introduced by the complementiser *ki*, which always precedes finite clauses.[39] An example is provided in (279):

(279)   *yo sab*   *ki*   [*use foy*   *Una*].
      1s know.NPST CMP 2   go.PST U.

---

[39] But see this chapter fn. 4

'I know that you've been to Una.'

Utterance complements in direct speech[40] are often also introduced by *ki*, as shown in (280a,b,c), although they may also occur without any overt complementiser (280d,e):

(280)   a.   rabbit *fal-o*   *ki*   [*bam*   *nɔs doy a*   *faze*   race].
           rabbit speak-PST CMP HORT 1p two IRR.NPST make.INF race
           'The rabbit said: "Come, the two of us will make a race".'

        b.   *aŋe*   *si ver*   *ad*   *diz-e*   *ki*   [*es*   *así̃ purtəgez*   *kẽ*
           someone if see.INF IRR.NPST say-INF CMP DEM$_p$ thus portuguese who
           fal]*?*
           speak.NPST
           'If someone sees, [they] will say: "Who speaks Portuguese like this?".'

        c.   *dəpəy dis*   *ki*   [*faz*   *nə,*   *d-use ɔm*   *vay*   *fik-a*
           then say.PST CMP make.NPST REQ of-2   man go.NPST become-INF
           mufin d-es].
           sulky of-DEM$_p$
           'Then [they] said: "Do it, [or] your man will become sulky because of that".'

        d.   *Fabian tə*   *fal-a*   [*mĩ*   *irmãw te*   *nə*   *d-el*
           F.   IPFV.NPST speak-INF 1s.OBL brother EXS.NPST REQ of-3s
           nom   *ε*   *Flavius*].
           name COP$_i$.NPST F.
           'Fabian is saying: "I have a brother, you see, his name is Flavius".'

        e.   *makak vey*   *i*   crocodile *dis*   [*d-use kurəsãw ɔn*   *te?*]
           monkey come.PST and crocodile say.PST of-2   heart   where EXS.NPST
           'The monkey came and the crocodile said: "Where is your heart?".'

On rare occasions, some of the older speakers (arguably under the influence of standard Portuguese) use indirect speech in these complements (281):

(281)   *εl dis*   *ikəl*   *di*   *a*   *mĩ ki*   [*el t-iŋ*   *vi-d*   *ẽ kaz i*   *el*
       3f say.PST DEM$_d$ day DAT me CMP 3s IPFV-PST come-PTCP in house and 3s
       kεr   *tum-a*   interview].
       want.NPST take-INF interview

---

[40]Cristofaro (2005b) debates whether or not direct speech constructions of this type should be interpreted as complements of utterance verbs. I treat them as complements here in the interest of simplicity, although I do not want to overemphasize any such claim.

'The other day, she told me: "he came to the house and he wants to make an interview".'

Indirect content questions are syntactically similar to direct interrogatives. (282) contrasts a direct interrogative (282a) with its indirect counterpart (282b):

(282) a. *use uki kɛr?*
           2s   what want.NPST
           'What do you want?'

       b. *dig   a   mĩ   [use uki   kɛr]*.
           tell.NPST DAT 1s.OBL 2s   what want.NPST
           'Tell me what you want.'

As demonstrated in (283), indirect polar questions are introduced by the particle *si* 'if', which is also a conditional subordinator. They are therefore similar in form to conditional clauses (see section 7.10.2):

(283) a. *use dig   nə   [si tə   fik-a   med də cockroach]*.
           2   say.NPST REQ if IPFV.NPST become-INF fear of cockroach
           'Do tell me whether you're afraid of cockroaches.'

       b. *ali   ve   [ũ   dɛmɔn si sa-i]*.
           there see.NPST one ghost   if leave-INF
           'Watch out if a ghost will come out.'

There are various examples of such complements, in which an indefinite, interrogative-like element simultaneously indicates subordination and occupies a constituent slot within the subordinate clause. It is not always tenable to interpret these as instances of reported interrogatives, but rather as indefinite complement clauses - identical to adverbial clauses, but in complement position. Some examples are given in (284):

(284) a. *yo nã   sab   [ɔn   yo pos   ikəl   papay ki   use de-w*
           1s NEG_cl know.NPST where 1s put.PST DEM_d papaya CMP 2   give-PST
           a   mĩ].
           DAT 1s.OBL
           'I don't know where I put the papayas that you gave me.'

       b. *pɔd   pɛrsɛb-e   [kẽ fal   d-ikəl   manɛr]*.
           can.NPST understand-INF who speak.NPST of-DEM_d manner
           '[You] can understand whoever speaks like that.'

233

c. [el kwɔn dis]   yo nã   sab.
3s when say.PST 1s NEG*cl* know.NPST
'I don't know when he said [it].'

So far, all of the complement clauses discussed have been balanced (in the sense of Stassen 1985), given that they still admit the full range of TAM marking as an independent clause. Let us now turn to instances of deranked complement clauses. Some verbs, known as *catenative verbs*, take a nonfinite complement clause without any overt subordinator. These include for instance *kere* 'to want', *mãda* 'to command' or *kaba* 'to finish', but the status of the complement clauses involved is not always entirely the same. One group of catenatives take a nonfinite complement clause only in the event that the subject of both the main clause and complement clause is coreferential[41]; otherwise, they take a finite complement clause flagged with an overt subordinator. The desiderative verb *kere* is one such catenative - for an alternative use of *kere* as an auxiliary, see (286) and section 7.1.2.4. This is demonstrated in (285). Sentence (285c) clarifies the monotransitive nature of the verb *kere*, which also takes nominal complements:

(285) a. yo nã   kɛr   [kõt-a], dig   a   el.
           1s NEG*cl* want.NPST tell-INF say.NPST DAT 3s
           'I don't want to tell, ask him.'

       b. yo kɛr   ki [use vay   imɔr].
           1s want.NPST CMP 2   go.NPST away
           'I want you to go away.'

       c. el kɛr   ũ   computer kɛr.
           3s want.NPST one computer want.NPST
           'He wants a computer.'

Another of these is the verb *sabe* 'to know'; as exemplified in (286a), its catenative use is also restricted to cases of subject coreference, and otherwise it may take as complement a finite *ki* clause (286b) or an indirect interrogative (286c) - see also the examples in (279). Furthermore, *sabe* also admits phrasal complements, as shown in (286d):

(286) a. yo sab   [faz-e   tud koyz].
           1s know.NPST make-INF all thing
           'I can make all sorts of things. [Lit. 'I know to make all things.']'

       b. yo sab-iŋ   ki [el tə   fal-a   Liza mem].
           1s know-PST CMP 3s IPFV.NPST say-INF L.   EMPH

---

[41]See, for example, Haspelmath (2005b) for a similar account of the complements of desiderative verbs from a crosslinguistic perspective.

'I knew that he would say "Liza" only.'

c. *kẽ   sab   [elz kwɔn / si a   v-i]?*
who know.NPST 3p   when   if IRR.NPST come-INF
'Who knows when/whether they'll come?'

d. *el sab   mariti.*
3s know.NPST Marathi
'He speaks [Lit. 'knows'] Marathi.'

Catenative verbs are often responsible for expressing certain modal and aspectual categories. Notice how the catenative use of *sabe* in (286a) expresses (mental) ability, which can also be aprehended from a construction such as (286d). Clear examples of this are the verbs *(a)kaba* 'to finish' and *kuməsa* 'to begin', which indicate inceptive and completive aspect respectively. There is a crucial difference between these verbs and *kere* 'to want' and *sabe* 'to know', for instance, in that *(a)kaba* and *kuməsa* do not take a complement clause in which the subject is not coreferential with that of the main clause. Let me first clarify that there is an intransitive use of *(a)kaba/kuməsa*, in which the subjects are undergoers rather than agents. This is exemplified in (287):

(287) *istɔr kuməs-o   / kab-o.*
story begin-PST   finish-PST
'The story began/ended.'

Transitive *(a)kaba/kuməsa* are rather different, as they can take agentive subjects. This is demonstrated in (288) for *(a)kaba*. in which the verb takes a complement NP:

(288) *nɔs əd   kab-a   es   tud.*
1p IRR.NPST finish-INF DEM$_p$ all
'We will finish all of that.'

Their occurrence as catenative verbs is much more common, though, in which case they characterise the action in terms of what has been called *phasal aspects* (e.g. Payne 1997), i.e. aspectual categories which refer to different phases of the action. *Kuməsa* marks inceptive aspect:

(289) a. *Astle vey   ku   ũ   mãduk, pos   isim   də Conchita, Conchita*
A.   come.PST COM one frog   put.PST on.top of C.   C.
kuməs-o [fik-a   med].
begin-PST become-INF fear
'Astle came along with a frog, he put it on top of Conchita, Conchita started getting frightened.'

b. *kwɔn muz vey jə kuməs-o [leŭt-a say]*.
   when music come.PST immediately begin-PST lift-INF skirt
   'When the music started, [she] immediately started lifting [her] skirt.'

The catenative use of *(a)kaba* indicates the end of an action, therefore marking completive aspect. This is demonstrated in (290):

(290) a. *nɔs kab-o [mar-a kaz]*.
          1p finish-PST build-INF house
          'We finished building the house.'

   b. *yo dis use akab [kum-e] lɔg i vay imɔr*.
      1s say.PST 2s finish.NPST eat-INF immediately and go.NPST away
      'I said: "Finish eating quickly and go away".'

In the continuum linking auxiliary verbs to catenatives (the scale of clausal integration discussed in 7.1.2), *(a)kaba* and *kuməsa* fall away from either of the extremes, but it is clear that they are more semantically prominent than prototypical auxiliaries; the occurrence of imperative *akab* in (290b) testifies to just that.

There is yet another group of catenative verbs which correspond to three-place predicates, in which the beneficiary argument of the main clause is coreferential with the agent of the complement clause. These include the verbs *dixa* 'let', *mãda* 'to command', *ajuda* 'to help' and a specific use of *faze* 'to make'.

The verb *dixa* 'let', the catenative use of which has already been discussed in section 7.1.2, indicates permission. The fact that the beneficiary of *dixa* is the subject of the complement clause, and also that the argument is expressed only once, motivates some variation: this argument may be treated either as an indirect object - recognisable through the Dative marking of the argument in (291a)[42] - or as a subject - for instance the unmarked occurrence of *el* in (291b). Arguments which are unimportant or easily retrievable from (discourse) context may be ellipted, to the extent that none at all is expressed; this is exemplified in (291c):

(291) a. *dex [[uv-i] pə Hugo]*.
          let.NPST listen-INF DAT H.
          'Let Hugo listen.'

   a. *nã dex [el [fuj-i]]*.
      NEG*cl* let.NPST 3s flee-INF

---

[42] Example (291a) is potentially ambiguous. The dative-marked argument *Hugo* could equally well be the (indirect) object of *dex* (in which case it would mean 'Let Hugo listen [to X]') or as the object of *uvi* (which would mean 'Let [X] listen to Hugo'). The correct interpretation is retrieved from context. See (292b) below for a comparable case of ambiguity.

'Don't let it run away.'

b. *(a)tɛ ki   nãw   cheg-a   ɔr,   dex   [fik-a].*
   until CMP NEG*cl* arrive-INF hour let.NPST stay-INF
   'Until the time comes, let it be.'

The catenative use of both *faze* 'to make' and *mãda* 'to command' (also 'send') produces causative constructions, with the distinction that the latter implies that the subject of the main clause has some sort of authority over the subject of the complement clause. Once again, there is variation as to whether the argument shared by the two clauses is treated as an indirect object or a subject. (292) exemplifies the use of *mãda*. Notice that, in (292a), only the subject of the main clause and the object of the complement clause are expressed, which leaves out the shared argument; the syntactic role of *elz* in (292b) could be ambiguous, but this ambiguity is resolved through context:

(292) a. *es   chef Rujɛr mãd-o   [da].*
          DEM*p* chief R.   command-PST give.INF
          'Chief Rogério ordered [them] to give this [to me].'

       b. *nɔs tə   mãd-a   [[faz-e   kim-a]   a   elz].*
          1p IPFV.NPST command-INF make-INF burn-INF DAT 3s
          'We make them roast [it].'

The causative semantics of catenative *faze* are clear from the example in (293):

(293) *yo fez   [[fuj-i]   pə   peacock].*
       1s make.PST flee-INF DAT peacock
       'I made the peacock run away.'

The corpus contains only one instance of the verb *da* 'to give' commanding a nonfinite complement clause. The interpretation seems to be similar to that of the causative constructions discussed above, coupled with a notion of physical transfer:

(294) *es   nã   pɔd   da   [[kum-e] pə   tud jẽt   də fɔr],*
       DEM*p* NEG*cl* can.NPST give.PST eat-INF DAT all people of outside
       ε   kar.
       COP*i*.NPST expensive
       '[We] can't give this for outsiders to eat, it's expensive.'

Finally, the use of *ajuda* 'to help' as a catenative verb is exemplified in (295):

237

(295) *yo ajud-o*   [*mama*   [*faz-e*   *bibik*]].
       1s help-PST auntie/granny make-INF bebinca
       'I helped autie/granny make bebinca.'

7.10.2 Adverbial clauses

Adverbial clauses are subordinate clauses which do not fill in a complement position, i.e. whose function is similar to that of adverbials. They are, in a sense, secondary to the predication, and yet they are responsible for contextualising the proposition and establishing logical relations of condition or purpose, among others.

Adverbial subordinators are often equivalent to relative or interrogative pro-forms. Clauses with a manner adverbial function are introduced by *kom/kufɔr* 'like', as shown in (296) - see also sections komprep and 7.11.2:

(296)   a. *nɔs fal*   *ikəl*   *mem*   *ẽ kaz*   *purtəgez*   [*kom nɔs fal*
            1p speak.NPST DEM$_d$ EMPH in house portuguese as   1p speak.NPST
            tud di].
            all day
            'We speak Portuguese at home like we do every day.'

        b. *asĩ* *tə*   *aprəs-e*   [*kufɔr ki*   *ŭ* speaker   *te*   *aki*].
            thus IPFV.NPST seem-INF   like   CMP one loudspeaker EXS.NPST here
            'It seems/sounds as though there was a loudspeaker here.'

Time adverbial clauses typically involve the subordinator *kwɔn* 'where', as shown in (297):

(297)   a. [*iskur nɔs kwɔn tə*   *vay*], *Conchita tə*   *fala*   *nã*
            dark 1p when IPFV.NPST go.INF C.   IPFV.NPST say.INF NEG$_{cl}$
            kɛ   *vay,*   *nã*   *kɛ*   *vay.*
            want.NPST go.INF NEG$_{cl}$ want.NPST go.INF
            'Whenever we go into the dark, Conchita says: "I don't want to go, I don't want to go".'

        b. [*kwɔn muz*   *vey*]   *jə*   *kuməs-o*   *leŭt-a*   *say.*
            when music come.PST already begin-PST lift-INF skirt
            'When the music started, [she] immediately started lifting [her] skirt.'

        c. *uvi,* [*Jizuz kwɔn mure-w*] *use foy*   visit?
            hear Jesus when die-PST 2   go.PST visit
            'Listen, when Jesus died did you go and visit [him]?'

238

In addition, time adverbial clauses can also be introduced by temporal prepositions such as *ãt* 'before' or *dəpəy* 'after', which must however be followed by *də* + Nonfinite verb form (if the subject of the subordinate clause is coreferential with the subject of the main clause) or *ki* + Finite verb form (if the subject of the subordinate clause and the main clause are distinct or if, although coreferential, they are repeated). Some examples are given in (298) - see also section 7.4.10:

(298) a. [*ãt   də mat-a   buf*]   *nã   vĕd   pɛl.*
           before of kill-INF buffalo NEG*cl* sell.NPST skin
           'Do not sell the skin before killing the buffalo.'

       b. *nɔs vay   durm-i   [dəpəy də tum-a   baŋ].*
           1s go.NPST sleep-INF after of take-INF bath
           'We go to sleep after we take a bath.'

       c. [*dəpəy ki   use foy*],   *chuv kuməs-o   ka-i.*
           after CMP 2   go.PST rain begin-PST fall-INF
           'After you left, it started raining.'

       d. *ikəl   pad   ki   t-iŋ   aki* [*ãt   ki   es   pad   chig-o*].
           DEM*d* priest REL EXS-PST here before CMP DEM*p* priest arrive-PST
           'The priest who was here before this priest arrived.'

Occasionally, the verb forms of temporal subordinate clauses of the latter type are nonfinite, as demonstrated in (299). Such constructions are quite rare.

(299) [*(a)tɛ ki   nã   cheg-a   ɔr*], *dex   fik-a.*
       until CMP NEG*cl* arrive-INF hour let.NPST stay-INF
       'Until the time comes, let it be.'

The combination of the adverb *lɔg* 'immediately' and *ki* results in a complex meaning 'as soon as', which can command a full clause as in (300):

(300) [*lɔg   ki   chig-o*]   *kums-o   traby-a.*
       immediately CMP arrive-PST begin-PST work-INF
       'As soon as [he] arrived, [he] started working.'

Spatial adverbial clauses are somewhate rare in the corpus, certainly so when compared to temporal adverbial clauses. The ones which do occur are introduced by *ɔn* but often preceded by a generic locative adverb. This is exemplified in (301)[43]:

---

[43] *Pɔrt də kãp* is the name of one of the ancient gates of the city wall, from Ptg. *Porta do campo* 'door of the fields' - see section 8.6.4

(301) *Pɔrt də kãp ε   ali [ɔn   te   ikəl   theatre].*
       Pɔrt də kãp COP$_i$.NPST there where EXS.NPST DEM$_d$ theatre
       'The *Pɔrt də kãp* is there where the theatre is.'

Reason clauses are typically introduced by *pərki/purke* 'because'. The subordinate clause in such cases is finite, as exemplified in (302):

(302) a. *Conchita ε   mεdroz [pərki əl tə   fik-a   də tud*
           C.   COP$_i$.NPST fearful because 3s IPFV.NPST become-INF of all
           koyz med].
           thing fear
           'Conchita is very easily frightened because she is scared of everything.'

       b. *õt   noyt nɔs t-iŋ   kãsad [purke nɔs t-iŋ   nad-a*
           yesterday night 1p COP$_s$-PST tired because 1p IPFV-PST swim-INF
           ĩter   di].
           whole day
           'Last night we were tired because we had been swimming the whole day.'

Alternatively, the preposition *(pu)kawz* may be employed, as in (303), provided that it is followed by *ki* (see section 7.4.5):

(303) *nɔs nã   foy   fεs [kawz   ki   nɔs t-iŋ   bastãt trabay].*
       1p NEG$_{cl}$ go.PST party because CMP 1p have-PST much work
       'We did not go to the party because [Lit. 'for the reason that'] we had a lot of work (to do).'

The subordinator *pu* (sometimes realised *pə*, but fundamentally distinct from the dative marker) introduces adverbial clauses of purpose. This subordinator commands strictly nonfinite (i.e. deranked) subordinate clauses, as shown in (304)[44]:

(304) a. *el t-iŋ   vay   nə   ũ   jungle [pu   traz-e   koys, aros tud].*
           3s IPFV-PST go.INF LOC one jungle PURP bring-INF thing rice SIML

---

[44]The semantic connection between purpose and reason adverbials/adverbial clauses is evident. Although DIP makes a formal distinction between the two, notice that the purpose clauses in (304) can easily be replaced with equivalent reason clauses. The following is a possible alternative to (304b) which was also produced during elicitation:

(1) *el sẽp   dəpəy də jãtar ad   vay   kaz [purke el a   ker-e   drum-i].*
     3s always after of lunch IRR.NPST go.INF house because 3s IRR.NPST want-INF sleep-INF
     'He always goes home after lunch because he needs to sleep.'

'He went into a jungle to bring some things, rice and all that.'

b. *el sẽp dəpəy də jãtar ad vay kaz [pu drum-i].*
3s always after of lunch IRR.NPST go.INF house PURP sleep-INF
'He always goes home after lunch to sleep.'

c. *trag ikəl [pu uv-i].*
bring.NPST DEM*d* PURP listen-INF
'Bring that [computer] so we can listen.'

Conditional constructions consist of two clauses with interdependent truth values, one of which is a subordinate conditional clause (known as *protasis*) introduced by the subordinator *si* 'if'. The verbal head of the main clause (i.e., the *apodosis*) is often modified by the irrealis marker while the verb of the conditional clause occurs in its infinitival form[45], as shown in the following example. In both sentences, the subject of both clauses is the same, and yet it is repeated in (305b):

(305)   a. [*officer si fik-a sabe-n*] *nã ad gost-a.*
officer if become-INF know-PTCP NEG*cl* IRR.NPST like-INF
'If the officer finds out, he won't like it.'

   b. [elz si kum-e pod kumid], elz a fik-a duẽt.
3p if eat-INF rotten food 3p IRR.NPST become-INF sick
'If they eat bad food, they will be sick.'

   c. [si cham-a], mama ad v-i.
if call-INF auntie/granny IRR.NPST come-INF
'If you call, auntie/granny will come.'

Occasionally, the irrealis marker may attach to the verb in the protasis instead, which seems to express a hypothetical rather than factual condition. In (306a), the verb in the apodosis is imperfective because it is a generic statement, but in (306a) the head of both clauses is irrealis-marked:

(306)   a. [*si yo a beb-e beb-e*] *tə durm-i.*
if 1s IRR.NPST drink-INF drink-INF IPFV.NPST sleep-INF
'If I were to drink a lot, I would [always] fall asleep.'

---

[45] In the speech of certain adult speakers with a more solid knowledge of SP, a Portuguese Subjunctive verb form occupies the head of the conditional clause, see *tivɐr* in (308) below. Subjunctive forms, quite rare in the corpus, may actually be instances of code-switching towards SP rather than part of the DIP verbal paradigm.

b. [*si chuv n-əd   chuv-e*], *nɔs a   vay   nə   bazar   də*
   if rain NEG-IRR.NPST rain-INF 1p IRR.NPST go.INF LOC market of
   pex.
   fish
   'If it doesn't rain, we will go to the fish market.'

As mentioned in 7.1.2.2 above, the occurrence of the Past Irrealis marker *vidi* within a Conditional construction expresses counterfactual mood:

(307)   [battery *aki si te*],   *chalu   vidi   fik-a.*
        battery here if EXS.INF functioning IRR.PST become-INF
        'If there were a battery here, (it) would be functioning.'

Occasionally, conditional constructions trigger the occurrence of the adverb *ãtã* 'then' introducing the apodosis. Although this is not required or predominant, there is a strong tendency for this construction to be used in such constructions.[46] An example is given in (308):

(308)   *agɔr* [*si t-ivɛr   nɔven   də nɔsior   də mɔrt, nɔven   də nɔsior   də*
        now if EXS-SBJV novena of our.lady of death novena of our.lady of
        imakulad], *ãtã ɛ   nə   igrej.*
        immaculate then COP$_i$.NPST LOC church
        'Now if there is the novena of Our Lady of Death, the novena of Our Lady of Immaculate [Conception], then [mass] takes place in the church.'

7.10.3   Relative clauses

Relative clauses are subordinate clauses which occur within the NP and serve to specify or circumscribe the reference of the nominal head - all naturally-occurring relative clauses

---

[46]Conditionals are not the only subordinate clauses to trigger the occurrence of *ãtã* in their matrix clauses, but also time adverbial clauses with a hypothetical reading (1a) and suppositions (1b). The crucial element common to all these constructions seems to be *hypothetical semantics*:

(1)   a. *kwɔn v-ir   aɲe   peso də fɔr,   ãtã nɔs fal   pok may* melhor *asĩ.*
         when come-INF someone person of outside then 1p speak.NPST little more *melhor* thus
         'When some outsider comes, then we speak a little bit better like this.'

    b. suppose *yo agɔr tə   kõt-a   ũ   istɔr ku   kwɔlkɛr amig də mĩ*   [...] *ãtã*
         suppose 1s now IPFV.NPST tell-INF one story with any   friend of 1s.OBL   then
         yo fal   uki ɛ,   kɔm vay   Anju, te   bõ?
         1s speak.NPST what COP$_i$.NPST how go.NPST A.   COP$_s$.NPST good
         'Suppose I am now chatting to some friend of mine [...] then I say: "what is it, how are you Anju, are you alright?".'

Notice that the speaker who produced these sentences is a middle-aged woman whose speech features a number of acrolectal traits, such as for example the use of the SP adjectival form *melhor* 'better'.

in the corpus are of the restrictive type. The relativised noun is marked within the relative clause either by the relative particle *ki* or a set of relative pro-forms (see 6.6.4) which are isomorphic with interrogative pro-forms. A token count reveals that relative pronouns are much less frequent than the relative particle, and yet they are sometimes required because the relative particle cannot be case-marked (in the case of DIP, marked with a case-assigning preposition). Notice the example in (309), in which the relativised element is dative-marked. The particle *ki* is excluded from such a context, being replaced with the human relative pronoun *kẽ* 'who(m)'[47]:

(309) *ikəl   ɔm [a   kẽ/*ki use atər-o].*
       DEM_d man DAT REL   2   push-PST
       'That man whom you pushed.'

On the other hand, whenever case-marking is not applicable, *ki* and one or more relative pronouns may be entirely interchangeable. (310a) shows the equivalence of the relative particle and the locative relative pronoun *ɔn* 'where', while in (310b) the competing forms are the relative particle and two equivalent manner relative pronouns *kom* and *kufɔr* 'how' (see also 7.11.2):

(310) a. *ikəl   kaz [ɔn/ki use fik].*
           DEM_d house REL   2   stay.NPST
           'The house where you live/that you live in.'

       b. *maner [kom/kufɔr/ki nɔs fal   ẽ kaz].*
           manner REL   2p speak.NPST in house
           'The way how/that we speak at home.'

The relative marker may occur either in the beginning of the relative clause (311a,b) or in focus position, immediately preceding the VP (311c).[48]

---

[47]Very rarely (and particularly in unnatural circumstances such as elicitation), speakers resolve the conflict between the primacy of the relative particle and its block on case-marking with a resumptive pronoun which diverts case-marking from the relative particle. Although this construction is not ruled out, it must be stressed that this is not at all the standard relativisation strategy in DIP. The resumptive-pronoun equivalent to (309) is as follows:

(1) *ikəl   ɔm ki   use atər-o   a   el.*
     DEM_d man REL 2   push-PST DAT 3s
     'That man whom you pushed.'

[48]The same possibility was also observed for the *Norteiro* (Bombay-area) varieties of Indo-Portuguese by Dalgado (1906), who provides the following examples: *tud eu qui pedí* 'all that I ask'; *êll mezinh dentr qui tá pusá* 'what(ever) he adds to the potion'; *[ó]ss qui já deu anel* 'the ring that you gave'.

(311) a. *ikəl   di [ki   yo vi-w   a   use] t-iŋ   muyt sɔl.*
DEM$_d$ day REL 1s see-PST DAT 2   EXS-PST much sun
'The day that I saw you was very sunny.'

    b. *yo foy   n-ikəl   igrej [ɔn yo kaz-o].*
1s go.PST in-DEM$_d$ church REL 1s marry-PST
'I went back to the church where I got married.'

    c. *frãgiŋ [yo ki   kõpr-o]   ɛ   bõ.*
chick 1   REL buy-PST COP$_i$.NPST good
'The chicks that I bought are good.'

If the verb of the relative clause occurs in an atypical position, the relative marker is likely to migrate with it. In (312), the relative particle *ki* occurs in immediate preverbal position even though the VP appears in sentence-final position:

(312) [*pɛrt də kapɛl ki   te*]   *sɔl   kaz?*
near of chapel REL EXS.NPST DEM$_d$ house
'The house which is next to the chapel?'

Headless NPs (in which a nominal head is assumed) may still govern a relative clause, as shown in (313):

(313) *tud [use ki   fal-a]   ɛ   vɛrdad.*
all 2s   REL say-INF COP$_i$.NPST true
'Everything that you say is true.'

The syntactic analysis of sentences such as the ones in (314) is somewhat unclear:

(314) a. [*kẽ chig-a   pimer] a   gaŋ-a prɛzẽt.*
who arrive-INF first   IRR.NPST win-INF present
'Whoever arrives first will receive a present.'

    b. [*yo uki fal-a] use ti   ki   faz-e.*
1s what say-INF 2s AUX.NPST CMP do-INF
'You must do whatever I say.'

The sequences *kẽ chiga pimer* and *yo uki fala* in (314), which occupy argument positions in these cases and therefore constitute complement clauses, might be interpreted either as headless (free) relative clauses or indirect interrogatives, given the formal overlap between interrogative pro-forms and relative pronouns. One indication which may help

resolve the question is the occurrence of *uki* in (314b), given that this element does not occur in regular relative clauses at all - see section 6.9.1. There are two possible interpretations to the fact: one can take this as evidence that such non-specific complement clauses make use of interrogative pro-forms rather than relative pronouns; alternatively, one may admit that the form *uki* does constitute a relative pronoun, but one which is prevented from co-occurring with an overt (specific) nominal head.

7.11   Comparison

Despite their logical similarity, there are some clear-cut syntactic distinctions between comparative constructions proper (e.g. *more/less X than Y*), which encode the comparison of inequality, and similative constructions (e.g. *as X as Z*). I will therefore discuss these two constructions separately.

7.11.1   Comparative constructions

In DIP, the prototypical comparative construction features a *comparee* NP, a predicate which defines the *parameter* of comparison, an *index marker* on the parameter (usually *may* 'more', and less frequently *men* 'less'), a *comparative particle ki* 'than' and a *də*-PP which establishes the *standard of comparison*. (315) provides examples of this type of construction:

(315)   a.   *ɛlifãt   ɛ   may fɔrt   ki   də nɔs.*
             elehant COP$_i$.NPST more strong COMP of 1p
             'Elephants are stronger than us.'

        b.   *galiŋ   ɛ   may barat ki   də karner.*
             chicken COP$_i$.NPST more cheap COMP of mutton
             'Chicken is cheaper than mutton.'

Comparatives with *men* 'less' are extremely rare. The assessment of inferiority can usually be expressed either with a negated *may*-comparative or by alternating logically opposite adjectives (such as *big* vs. *small*).

Having set out the structure of prototypical comparative constructions, it is now possible to understand instances which depart form it in one way or another. Firstly, it is clear that both the comparee and the standard of comparison need not be single entities, they can be clauses with (a) shared constituent(s) (cf. Dixon 2008). In (316), it is both the verb (*te* 'to have') and the object (*nɛt* 'grandchild') which are shared by the two clauses, and the focus of comparison is quantity:

(316)   *yo te   may nɛt   ki   d-ɛl.*
        1s have.NPST more grandchild COMP of-3f

'I have more grandchildren than her.'

In (317), the intransitive verb - *kure* 'to run' in (317a), and *sai* 'to depart' in (317b) - is the only constituent shared by the two clauses. These examples show that the property being compared need not be expressed by an adjective either. In both sentences, the parameter is adverbial because what is being compared is manner, in (317a), and time of departure, in (317b); notice that, in the latter, no index is indicated given that the adverb *pimer* 'earlier' is inherently comparative:

(317) a. *el te   kur-e   may dəfors ki   də nɔs.*
           3s IPFV.NPST run-INF more quickly COMP of 1p
           'He runs more quickly than us.'

       b. *yo a   sa-i   pimer ki   d-use.*
           1s FUT leave-INF earlier COMP of-2
           'I will set off first [lit. earlier than you].'

In other instances, it may well be that the verb and all participants are coreferential in the two clauses and only an adjunct constituent differs. Consider the sentence in (318):

(318) *yo gɔs   kum-e   mays buliŋ   nə   damãw.*
       1s like.PST eat-INF more croquette LOC Daman
       'I prefer eating croquettes in Daman.'

Context had previously made it clear that what was being compared in (318) was the pleasure of eating croquettes in Daman and the pleasure of eating croquettes in Diu, location being the only differing variable. This sentence also clarifies that the ellision of material aprehensible from context can be rather extreme.

Truncated comparative constructions occur whenever the standard of comparison indicates quantity. In such constructions, *may*, without any overt parameter of comparison, occur immediately before the standard of comparison. An example is given in (319):

(319) *dəpəy say   fɔr,   may də kwɔrẽt an jə   fik-o.*
       after leave.NPST outside more of fourty year already become-PST
       'And then [he] goes out, [this was] more than fourty years ago.'

The corpus contains an interesting case of a discontinuous comparative construction, in which the standard of comparison precedes the comparee NP, with the repetition of the index marker. This is given in (320):

(320) *may ki   də nɔs   kristãw, chɛf də damãw ɛ   may bõ.*
      more COMP of 1p.POSS christian chief of Daman COP$_i$.NPST more good
      'More than our Christians, the chief of Daman is better [than them].'

7.11.2 Similative constructions

Similative constructions are, logically, a subtype of comparative constructions which establish the similarity (rather than dissimilarity) of the comparee and the standard of comparison. As shown in (321), in basic similative constructions the index of comparison may or may not be marked on the predicate with the similative marker *tãt* 'as much as'; the standard of comparison is usually expressed as a PP introduced by the preposition *kom* 'like':

(321) *Aline ɛ   (tãt) alt kom irmã.*
      A.   COP$_i$.NPST as   tall like sister
      'Aline is as tall as [her] sister.'

The functions of *kom* somewhat overlap with the preposition *kufɔr* 'like', although the latter is preferred in sentences which demonstrate a similarity in manner (see section 7.4.6). Notice their alternative use in (322). In both these sentences, a similarity in manner is clearly the target of the similative construction[49]:

(322) a. *yo vay   kaz bri̱k-a   mi̱   saykəl kom/kufɔr dhoom.*
         1s go.NPST house play-INF 1s.POSS bicycle like   *dhoom*
         'I am going home to play with my bicycle like they do in *Dhoom*.'

      b. *use fal   / aprɛs   kom/kufɔr d-use mãy.*
         2   speak.NPST seem.NPST like   of-2   mother
         'You talk/look like your mother.'

Kom and *kufɔr* also qualify as subordinators in that they can introduce adverbial clauses of manner. This is clear from the sentence transcribed in (323) - see section 7.10.2:

(323) *nã   fal   asi̱, nãw, asi̱ kufɔr yo fal   ku   syor nãw.*
      NEG$_{cl}$ speak.NPST thus NEG$_p$ thus like   1s speak.NPST COM mister NEG$_{cl}$

---

[49]It must be clarified that the verb *aprəse* 'to seem/look' often commands an adjectival complement, as in the following sentence:

(1) *es   foreigner aprɛs   bunit.*
    DEM$_p$ foreigner seem.NPST beautiful
    'This foreigner looks beautiful.'

,

247

'[We] don't speak like this, no. Not the way I'm speaking to you, no.'

Just like relative pronouns, subordinating *kom* and *kufɔr* are constituents of the subordinate clause in (323). This interpretation is blocked if these elements and the actual subordinate clause are mediated by the complementiser *ki*. This is exemplified in (324) below:

(324) *asĩ tə   aprəs-e   kufɔr ki   ũ   speaker   te   aki.*
       thus IPFV.NPST seem-INF like CMP one loudspeaker EXS.NPST here
       'It seems/sounds as though there was a loudspeaker here.'

Kom and *kufɔr* in fact do occur as relative pronouns as well, as shown in (325):

(325) *yo gos ikəl   maner kom/kufɔr use fal   ku   kriãs.*
       1s like DEM$_d$ manner how   2   speak.NPST COM child
       'I like the way you speak to children.'

For further information, refer to section 7.10.3.

Chapter 8

Lexicon

This chapter explores some aspects of the DIP lexicon, not only in terms of its formative processes but also certain salient semantic and etymological characteristics. A discussion of the lexicon is, by necessity, only partial, and the problem is especially acute when dealing with languages spoken by multlingual communities in multilingual environments (see the introduction to chapter 5). Section 8.1 discusses the challenges of multilingualism for the delimitation of the lexicon of DIP as an object of study, and advances some practical solutions. Section 8.2 shows the extent to which various compositional processes (have) operate(d) in building the language's lexicon. The following sections are dedicated to a structured account of particular semantic fields (section 8.6) and etymological sources (section 8.7).

Unlike the previous descriptive chapters, I will not enforce a strict segmentation of synchronic and diachronic discussion here, as etymological information is best interpreted when presented alongside the present-day lexeme(s). In this chapter, etyma - whenever identifiable - are given in square brackets, preceded by an abbreviated indication of the language of origin. Portuguese and Gujarati being the major players in the formation of DIP, and the languages which have for the longest period played a role in the linguistic environment of the island, it is not surprising that these two languages are well-represented as sources of lexical material for DIP, as is English. The Konkani element (section 8.7.2) is possibly more surprising, but very revealing. One should resist over-interpreting the relative proportions of their contributions, however, for the reasons enunciated in 8.1.

8.1 Imposing boundaries on the lexicon

Deciding on whether a particular recorded word should be regarded as a component of the DIP lexicon is not only difficult but also highly dependent on one's concept of *lexicon*. In the case of DIP, given the high level of multilingualism of its speakers, an approach which recognises the fact that speakers avail of a complex and fluid set of lexemes is powerful in accounting for the high levels of lexical variation. On the other hand, it reveals the difficulty of imposing boundaries on the language's lexicon by any objective means.

8.1.1 The lexicon as a pool of lexemes

If a particular speech community is highly multilingual, with high levels of proficiency in more than one language, it is fair to say that any speaker avails of a 'pool' of lexical possibilities[1] containing competing forms with (partially) overlapping semantics. A crucial factor in this scenario is that the option for any of the competing forms does not compromise the effectiveness of the communicative act. The speakers may be able to assign each of the competing lexemes to a particular 'lect' (language, sociolect, dialect, idiolect, or any such abstraction), which is an important structuring factor when it comes to select one of the forms for a particular utterance. The expected knowledge of the interlocutor(s) is one of several constraining factors, but if all participants are expected to share more than one of the competing forms it will only be of secondary importance. The challenge this poses to the descriptive effort is that, while in a sense every lexeme known to the multilingual speaker could be said to constitute that speaker's lexicon, one is expected to classify this lexical universe into 'language' categories. This is reminiscent of the difficulty of establishing cut-off points in the cline between a(n occasional) borrowing and a fully nativised lexeme, either at the abstract level of communal language or the individual.[2]

For this specific description, the challenge is to decide which lexemes used by DIP speakers - be them ultimately of Gujarati, Portuguese, English or any other origin - should be included in a hypothetical 'lexicon of Diu Indo-Portuguese'. It is very clear, in this case, that speakers are aware of various competing forms for particular referents, and that the selection of one or the other is influenced by such factors as the expected knowledge of the interlocutor[3], etymological considerations, normative pressure and purism, statistics of exposure to the various forms, the target language, personal preferences, family preferences, social preferences, or even chance.

8.1.2 Practical criteria

A multilingual lexicon is highly fluid, and so are the selections effected by a given speaker. The following criteria were designed to limit somewhat the lexicon one might consider central to DIP, but I do not claim these to be universally applicable or free from arbitrariness.

---

[1] The term 'pool' is intentional here, as it links to an ecological conception of language creation and change (see e.g. Mufwene 2001) according to which speakers in language formation/acquisition are faced with a diverse pool of overlapping or even contradictory options ('features', be them syntactic, phonological, lexical or otherwise) from which they must select the ones to attribute dominance to. This concept is equally enlightening when applied synchronically, as it captures the diversity of linguistic input and knowledge available to any speaker - even those considered 'monolingual' - at any given time.

[2] The gradience of these facts also poses a practical problem to the definition of bilingualism given that, as Lehiste (1988:1) puts it, 'the theoretical limits to bilingualism might be drawn to encompass the range between the person who uses one nonintegrated loanword and the so-called perfect bilingual who can pass for a monolingual in more than one language'.

[3] This factor is one of components of what has been termed the *observer's paradox* in the literature on linguistic inquiry and documentation (see Labov 1966, 1972, Milroy 1987), i.e. the fact that any researcher inescapably influences the output of their collaborators, and therefore cannot expect to collect completely unconstrained speech. The speech data imagined to be the most reliable is that between interlocutors with maximal shared linguistic knowledge, which is precisely that which imposes less clear constraints on the selection of competing lexemes.

The words in the group thus constrained are the ones I base most of the present chapter on. It is also interesting to notice which semantic fields, if any, might be more prone to the use of lexemes from a particular source, which means I have also taken count of the occurrence of words whch, for the various reasons enunciated below, have not made it to the immediate lexicon of DIP.

All words which occur repeatedly in the corpus without any perceivable competitors were included in the lexicon. So were similar lexemes occurring only once, which were particularly common as the product of lexical elicitation, unless the collaborators specifically tagged them as 'foreign'.[4] For instance, when eliciting the word for 'tweezers', my collaborators indicated *chipia/chipio* (Guj. *chipyo* 'tweezers') with no alternative form. Considering that the word does not occur anywhere else in the corpus, it has been excluded from the DIP lexicon. In elicitation, there was a particularly clear bias towards words perceived as being of Portuguese origin in detriment of all others. Rather than interpreting this as a disparaging proof of speech constrain, it is important to recognise the cause of this preference - explained in detail in 2.2.1 - and treat it as an essential feature of the language.

Whenever two or more words are perceived as competitors because they have very close semantics, the following criteria were used. After compiling the occurrences of all competing lexemes within sentences with DIP as the matrix language, if the occurrence of one or more of these is strikingly low when compared to their competitors, they are considered non-dominant and therefore excluded from the lexicon. This was the case with the competing adjectives *mɛdroz* and *beɖka*, of Portuguese and Gujarati origin respectively, both of which describe someone as 'prone to fear' (see also section 2.2.3). In the corpus, *mɛdroz* occurs 12 times while *beɖka* occurs only twice, so the first form has been included in the lexicon but not the second one. A similar situation obtains when we compare the occurrence of *igrej* and *church*, of Portuguese and English origin, meaning 'church', with the first form emerging as highly dominant.

Pairs in which none of the variants can be said to be clearly dominant include, for instance, *pray* [from Ptg. *praia* 'beach'] and *bãdər* [from Guj. *bunder* 'port'], both of which are used toponymically to refer to a particular coastal stretch inside Diu Town (cf. 8.6.4).

Phonological considerations may also be taken into account, but with caution. The lexical elements of Portuguese origin, for their numerical dominance and centrality are shaped according to basic phonological rules which may or may not be imposed onto lexical elements from other sources. Whenever this occurs (e.g. neutralising original retroflex phonemes, shifting stress onto the final syllable), one might be more inclined to consider such lexemes as a more or less permanent borrowing, accepting the premise that frequency of occurrence favours phonological adaptation (Haugen 1950, McMahon 1994). However, the absence of phonological adaptation does not necessarily signal distance from a 'core' lexicon of DIP; particularly so when not only the individual but also the community have very similar levels of fluency in both languages involved in the process (see Thomason and

---

[4]Admitting single occurrences into the lexicon, either those which surface only in elicitation or in free-flowing speech, is necessary in the interest of comprehensiveness, but the method begs the recognition of a caveat. If further data is at some point collected and analysed, or if the corpus were in any way different, competing forms might be uncovered which could even be proven dominant, and therefore the output could be different.

Kaufman 1988:33). On the other hand, it is not uncommon for a frequent non-Portuguese-derived lexeme to be variably made to fit the central phonological constraints of DIP - see the outset of chapter 5 for concrete examples and a fuller discussion of these issues.

The task of assigning particular lexemes to the core vocabulary of DIP may still profit from other kinds of considerations. Take the example of the relatively infrequent pair *chaknŭ* (presumably derived from the Gujarati verbal root *chək-* 'to taste') and *biting* (from English). Both of these words were recorded in Diu to refer to a 'communal snack, light meal' and they occur a similar number of times. However, the speakers were clear in assigning the option *chaknŭ* to Daman (IP) and *biting* to Diu.

The criteria mentioned here, and often a combination of various criteria, have guided the definition of what to consider the core lexicon of DIP, which, from the perspective of various possibilities competing for selection by the speaker, amounts to the words which are most likely to be associated with this language and its sociocultural space.

8.1.3 Macrofunctionality

In this description, I prefer the term *macrofunctionality*[5] over *polyfunctionality* (or *multifunctionality*) because the former does not presuppose any semantic and functional 'primitives', which run the risk of being modelled on third-party languages. The approach associated to the notion of macrofunctionality, most evidently defined in Gil (2004), does not measure the observed differences in scope of (near-)translation equivalents according to a pre-established (often Eurocentric) standard but instead posits less well-defined functions which need not be broken up into more specific subfunctions by all languages and/or in the same way:

> Unlike polyfunctionality, which presupposes a plurality of functions, *macrofunctionality* assumes a single function, which may or may not be viably decomposable into a set of constituent subfunctions (Gil 2004:372)

The DIP word *ku* is therefore a classical macrofunctional element in that it functions not only as a comitative and instrumental preposition (see section 7.4.7) but also an NP-coordinator (see section 7.9.1). It is interesting to notice that the grammatical functions this element conflates in DIP are also expressed by a single element in many languages of the world (Stassen 2000, 2005b, Haspelmath 2004).

On a less strictly grammatical sphere, the semantic value of the DIP verb *kere* (described in 7.1.2.4), simultaneously desiderative (i.e. 'to want'), obligatory (i.e. 'must') and

---

[5]By definition, homonymy is not interpreted as macrofunctionality or macroreferentiality. In DIP, word pairs are interpreted as homonyms if a) the words belong to clearly distinct paradigms (e.g. the noun *kaz* 'house' and the verbal form *kaz*, the Non-Past form of the verb *kaza* 'to marry'.), b) if one of the elements in the apparent homonymous pair is a reduction of a different form (e.g. the Irrealis marker *ad*, which in its frequent reduced form *a* might be mistaken for the Dative preposition *a*), or c) if the two elements derive from different etyma and are still interpreted by the speakers as such (e.g. the adjective *kar* 'expensive' and the noun *kar* 'meat', from Ptg. *carne* 'met', which in DIP only sometimes retains the nasal segment ans is realised *karn*).

predictive (i.e. 'shall'), need not be seen as composite just because it corresponds to various elements in a different language. In other words, we can think of macrofunctional elements 'in terms of 'semantic vagueness' instead of 'semantic ambiguity'' (Khanina 2008:843). Even though a particular element may be highly macrofunctional, this does not mean finer semantic distinctions are absent from the language. In this particular case, while *kere* marks obligation among other modal categories, DIP has a competing compound auxiliary *ti ki* which exclusively marks obligatory mood (see 7.1.2.1).

The DIP verb *fika* (Past form *fiko*, Non-Past form *fik*) is an interesting case of multi-functionality, given its wide semantic and functional scope. Part of the semantics of this verb, I would argue, could be subsumed under the notion of (LASTING) ACCOMPLISHMENT. It occurs commonly with an adjectival complement, usually indicating a change of state, close in meaning to 'to become':

(326) *pə   el fik-o   bẽ   duẽt nə.*
       DAT 3s *fika*-PST very ill   REQ
       'He fell seriously ill, you know.'

In other cases, one detects a more stative semantic value. In (327), the meaning of the verb is closer to 'to remain':

(327) *el nuk   nə   fik   kalad mem.*
       3s never NEG$_{cl}$ *fika*.NPST silent EMPH
       'He is never quiet.'

These various readings carry over to instances of *fika* with a nominal or participial complement:

(328) a. *el fik-o   ĩtrujad   i   el foy   kaz.*
           3s *fika*-PST deceived and 3s go.PST house
           'He felt [became] deceived and he went home.'

       b. *use tə   fik-a   med də lagərtix?*
           2   IPFV.NPST *fika*-INF fear of gecko
           'Are you scared of geckos?'

Used intransitively, *fika* has eventive semantics, close in meaning to 'to happen':

(329) *uki fik-o?   fog fik-o.*
       what *fika*-PST fire *fika*-PST
       'What happened? There was a fire.' [lit. A fire happened.]

With an argument which happens to be a quantified expression of time, this verb refers to a past stretch of time in relation to the moment of the utterance or any other standard, a function roughly equivalent to the English adverbs 'ago' and 'earlier':

(330)　*may də kwɔrẽt an   jə   fik-o.*
　　　more of forty   year already *fika*-PST
　　　'More than forty years ago.'

Governing a localising PP, *fika* variably means 'to dwell' (331a) and 'to stay' (331b):

(331)　a.　*ikəl   muyɛr fik   de mĩ   lad.*
　　　　　DEM$_d$ woman *fika*.NPST of 1s.POSS side
　　　　　'That woman lives next to me.'

　　　b.　*ali   nã   vay,   aki mem fik.*
　　　　　there NEG$_{cl}$ go.NPST here EMPH *fika*.NPST
　　　　　'Don't go there, stay right here.'

Finally, I would argue that referentiality may be approached within the framework of macrofunctionality. Languages do vary as to the number and nature of distinctions they admit to refer to what could be seen as referential continua, such as colours or body parts. One case in point concerns the DIP word *pɛ*, which refers to both 'foot' and 'leg' while Portuguese distinguishes *pé* 'foot' from *perna* 'leg'. This instance of apparent semantic extension, which ammounts to the annulment of a semantic distinction, is in fact calqued on Gujarati, in which *pəg* refers to both 'foot' and 'leg'. Similarly, the DIP word *mãw* refers to 'hand' and 'arm', where Gujarati uses *hath* for both 'hand' and 'arm' and Portuguese distinguishes *mão* 'hand' from *braço* 'arm'.

8.2　Derivation

Derivation is not widespread in DIP, but certain patterns can nonetheless be recognised. This section describes some productive derivational processes in this language, as well as the issues raised by what must be seen as remnants of (no longer productive) derivational instruments - for instance the reinterpretation of the Ptg. diminutive suffix as an indicator of gender (section 8.2.1).

8.2.1　The status of the suffix -*iŋ*

The suffix -*iŋ*, recognisable in a few DIP words, is clearly derived from Ptg. -*inho*/-*inha*, which in Portuguese are still productive diminutive suffixes attaching to both nominal and adjectival stems. In contrast, some DIP words displaying this sequence can be said to

have lexicalised - and reanalysed - the suffix. Consider for instance the case of the pairs *avo* 'grandfather' vs. *avziŋ* 'grandmother', and *nɛt* 'grandson' vs. *nitiŋ* 'granddaughter', in which the suffix does not mark the diminutive but rather feminine gender (see section 8.7.4).

In other cases, words ending in *-iŋ* still establish contrasting pairs in DIP in which these forms are perceived as smaller than the unmarked one. Such pairs include:

- *vak* 'cow'	*vakiŋ* 'calf'
- *rat* 'rat'	*ratiŋ* 'mouse'
- *pas* 'bird'	*pasriŋ* 'small bird'
- *kals* 'trousers'	*kalsiŋ* 'underpants'
- *bol* 'cake'	*buliŋ* 'biscuit, croquette'
- *pok* 'little'	*pokxiŋ* '(very) little'

Despite these pairs of words, the suffix is certainly no longer (widely) productive in DIP. Another indicator of the (nearly) complete lexicalisation of the suffix is the word *ladiŋ* 'chant' [from Ptg. *ladaínha* 'chant']. In its present form, and if the suffix were still interpreted as such, *ladiŋ* might be interpreted as the diminutive of *lad* 'side' [from Ptg. *lado* 'side']. Yet, these two words occur in very different contexts and do not seem to be interchangeable.

In the case of the word *agəriŋ/alguriŋ* 'earring(s)' (from Ptg. *argolinha* '(ear)ring', composed of *argola* 'ring' + the diminutive suffix), the suffix was formally retained in DIP but not the diminutive semantics, given that the base lexeme of the Ptg. form (*argola*) is not part of the DIP lexicon. Other words with a similar ending but no productive diminutive contrast include *galiŋ*[6] 'chicken', *saldiŋ* 'sardine', *kuziŋ* 'kitchen' and *subriŋ* 'nephew/niece'.

8.2.2 The suffix -*sãw*

The suffix -*sãw*, present in various words of Portuguese origin such as for example *pursəsãw* 'procession' [from Ptg. *procissão* 'procession'] and *eləsãw/eleysãw* 'election' [from Ptg. *eleição* 'election'], sometimes occurs in combinations unattested in Portuguese. One example is *kuziŋsãw*, in which the suffix attaches to the root *kuziŋ*-, present in the verb *kuziŋa* 'to cook' and the noun *kuziŋ* 'kitchen':

(332) *faz-e   trabay, kuziŋ-sãw.*
       make-INF work   cook-*sãw*
       'To do (all the) work, (all the) cooking.'

---

[6]The Portuguese etymon itself, *galinha* 'chicken', does not carry diminutive semantics, instead producing a gender distinction with *galo* 'rooster'. No diminutive semantics is evident from the Ptg. words *sobrinho/sobrinha* 'nephew/niece' and *cozinha* 'kitchen' either.

In tune with the function of the suffix elsewhere (including in Portuguese), -sãw has a nominalising function.

8.2.3 The suffix -*m*

The suffix -*m* is still used, in what can be described as acrolectal DIP, to derive ordinals from cardinal numerals. Its use and productivity are fully described in 8.4.2.

8.2.4 Crystallised suffixes

Whereas suffixes -*iŋ* and -*sãw*, described above, raise questions concerning their productivity in modern DIP, other etymological derivational suffixes appear entirely unproductive. One of these is the suffix -*er*, from the Ptg. suffix -*eiro*/*eira* used to derive nouns (semantically agents or producers) from nouns. In DIP, this former suffix features in the name of various professions (e.g. *ĩfermer* 'nurse', *kuziŋer* 'cook'), characteristics (e.g. *patrater* 'show-off' [from Classical Portuguese[7] *patarateiro* 'pretentious person']) and plants (e.g. *kuker* 'coconut tree', *jaker* 'jackfruit tree', *mãger* 'mango tree'). Because this suffix is no longer productive, speakers resort to an analytical rather than suffixal strategy when in need of deriving new fruit-tree names: e.g. *papay* 'papaya' > *arv də papay* 'papaya tree'.

Likewise, the etymological suffix -*ãw* (not to be mistaken for -*sãw*, see 8.2.2) - from the suffix -*ão* used in Portuguese as an augmentative - is not productive in DIP. Nonetheless, among the various DIP words ending in this sequence, a few were originally derived from augmented Portuguese. etyma, including *fakrãw* 'sabre' (from Ptg. *facalhão*, derived from Ptg. *faca* 'knife') and *limãw* 'lemon' (from Ptg. *limão*, related to *lima* 'lime').

8.2.5 Hypocoristics

Hypocoristics are defined as (often short) names derived from full proper names, used as an informal form of address. In DIP, typical hypocoristics involve the reduction of the full name and the application of the ending -*u*.[8] Some examples are given below - notice that the proper names are given in their official orthography, while the corresponding hypocoristic follows the DIP orthographic conventions set out in 5.5:

- Anita > Nitu
- Angelina > Ãju
- Fátima > Fatu
- Gilberto > Jibu

---

[7]I use the term *Classical Portuguese*, with Cuesta and Luz (1980) and Castro (1991), to refer to the historical phase of the language spanning the period between the 16th and the 18th centuries; the date signalling the start of the Classical period is taken by several scholars to be 1536, when Fernão de Oliveira's *Grammatica da lingoagem portuguesa* (the first grammar of the Portuguese language) was published. See also Silva (1989, 1991), Teyssier (2001) and Cardeira (2005).

[8]A similar strategy is reported for both Korlai IP and Daman IP in Clements and Koontz-Garboden (2002:213).

- Severina > Xɛvu
- Virgilo > Vaju

There is one example of such a hypocoristic which does not derive from a person's name but from an ethnic appelation: natives of Gujarat (i.e. Gujaratis) are often depreciatively known as *gajus*. Less common alternatives involve the use of an *-i* ending, as in:

- Conchita > Kõchi

8.3 Compounding

The DIP lexicon contains a few endocentric compounds consisting of two nouns (e.g. *bix-kabel* '[a type of pilose bug]', composed of *bix* 'bug' and *kabel* 'hair'; *bix-fog* 'firefly', composed with *fog* 'fire') or a noun and an adjective (e.g. *bix-baboz* 'slug', composed of *bix* 'bug' and *baboz*[9] 'slimy') - notice that, in the latter case, the adjective appears after the noun, which would be an unusual word order if this were a syntactic construct.

The peculiarity of these compounds is that they shun the use of the generic relator *də* (see also 7.4.3), which allows non-prototypical elements (like nouns) to modify a nominal head.[10] The availability of the generic relator strategy means that its products are by definition syntactic constructs rather than lexicalised compounds. That is the case in the following: *bazar də pex* 'fish market' (*bazar* 'market' + *pex* 'fish'), *rɛlɔj də pared* 'wall clock' (*rɛlɔj* 'watch/clock' + *pared* 'wall'), *bɔs/mɛr də vak* 'cow-dung' (*bɔs/mɛr* 'dung' + *vak* 'cow'), *mak də kustur* 'sewing-machine' (*mak* 'machine' + *kustur* 'sewing'), *kalãw də mat* 'earthen pot' (*kalãw* 'pot' + *mat* 'earthenware'), or *prat də vid* 'glass dish' (*prat* 'dish' + *vid* 'glass').

Some generic relator constructions, however, can be interpreted as lexicalised compounds if a) the meaning of the construct is more specific than could be expected from a run-of-the-mill attributive construct (e.g. *bix-də-tɛr* 'earthworm', from *bix* 'bug' and *tɛr* 'earth'), or b) one of the components does not occur outside that particular construction (such as *kariŋ* in *kariŋ də liŋ* 'thread cylinder'[11], with *liŋ* 'thread').

---

[9]The term *baboso/babosa* was current in the 16th-century Portuguese spoken in India; the 1592 *Tombo de Diu* (Pais 1592) refers to the *aloe vera* plant as *erva babosa* 'slimy plant'.

[10]Adjectives do not occur in a generic relator construction, which reveals an interesting characteristic of ethnonyms such as *ĩdian* 'Indian' or *purtəgez* 'Portuguese'. Whereas in several languages such words can function simultaneously as nouns and adjectives (e.g. Eng. *Indians* and *Indian women*), in DIP these are strictly nominal: although they are allowed in attributive position, they require the use of the generic relator: *ũ paspɔrt də ĩdian* 'an Indian passport', *jẽt də purtəgez* 'Portuguese people'.

[11]Portuguese has the compound *carrinho de linhas* 'thread cylinder', both components of which (*carrinho* 'cart' and *linha* 'thread') can stand on their own. It is likely that, despite the generic relator structure, DIP acquired this compound as a lexical unit.

8.4 Numbers

The DIP numeral paradigms (both cardinal and ordinal) show striking similarities with those of Portuguese, but also significant differences. As far as derivation is concerned, it is interesting to see that the composite character of the Ptg. cardinal numerals 16 through 19 is annulled in DIP (see section 8.4.1) and also that a hesitant derivative suffix is seen to derive ordinals from their corresponding cardinals (see 8.4.2). This section describes both series in turn, starting with the cardinal numbers.

8.4.1 Cardinals

Cardinals are organised on a decimal scale, and they function as follows: below twenty, each number has an underived label:[12]

1 - *ũ*	11 - *õz*
2 - *doy*	12 - *doz*
3 - *trey*	13 - *trez*
4 - *kwɔt*	14 - *katɔrz*
5 - *sĩk*	15 - *kĩz*
6 - *sey*	16 - *zasey*
7 - *sɛt*	17 - *zasɛt*
8 - *oyt*	18 - *zoyt*
9 - *nɔv*	19 - *zanɔv*
10 - *dez*	

After twenty (*vĩt*), numbers become composite following the pattern 'decade *i* number [1 through 9]'; e.g. *trĩt i doy* 'thirty-two', *setẽt i sɛt* 'seventy-seven'. The decade labels above ten are the following:

20 - *vĩt*
30 - *trĩt*
40 - *kwɔrẽt*
50 - *sikwẽt*
60 - *sɛsẽt*
70 - *sɛtẽt*
80 - *oytẽt*

---

[12] The similitude of the pairs 6/16, 7/17, 8/18 and 9/19 is surely not lost on the speakers of DIP, but it would probably be abusive to interpret these as instances of morphological derivation considering that a) the subset is very reduced; b) the derivation is by no means extensible to other pairs, i.e., not productive; and c) part of the etymological material from Portuguese has been lost (Pt. *dez*asseis 'sixteen', *dez*assete 'seventeen', *dez*oito 'eighteen', *dez*anove 'nineteen'), obscuring the fact that these forms were actually composed of the root for 'ten' (*dez*) plus the root of the numbers 6, 7, 8 and 9.

The word for a hundred is *sẽ*, which can optionally be modified by *ũ* 'one'; the related form *sẽt* is a general label for hundreds in combination with any subsequent number; e.g. *(ũ) sẽ* 'a hundred', *sẽt i ũ* 'a hundred and one'. Subsequent labels go as follows:[13]

200 - *duzẽt*
300 - *trey sẽt*
400 - *kwɔt sẽt*
500 - *kiŋẽt*
600 - *sey sẽt*
700 - *sɛt sẽt*
800 - *oyt sẽt*
900 - *nɔv sẽt*

The label for thousands is *mil*, which must be modified by a number; e.g. *ũ mil* 'a thousand', *trĩt i sĩk mil* 'thirty-five thousand'. Both *sẽt* and *mil* usually combine with *ũ* 'one' but, when further material follows, *ũ* may be left unexpressed; any other number in combination with *sẽt* or *mil* must be expressed; therefore, *ũ sẽt i kwɔrẽt i doy* or *sẽt i kwɔrẽt i doy* 'a hundred and fourty-two', but necessarily *oyt sẽt i kwɔrẽt i doy* 'eight hundred and fourty-two'; *ũ mil i trey sẽt i doz* or *mil (i) trey sẽt i doz* 'a thousand three hundred and twelve', but necessarily *sɛt mil (i) trey sẽt i doz* 'seven thousand three hundred and twelve'.

The most common label for a ten-thousand-strong unit is *lak* whereas a hundred-thousand-unit is known as a *kror*[14], both of which can be quantified with resort to numerals; e.g. *nɔv lak* 'nine lakh [i.e. ninety thousand]', *doy kror* 'two crore [i.e. two hundred thousand]'.

8.4.2 Ordinals

In comparison with the cardinal numerals, the ordinal series is rather unproductive and uncommon. The most frequent ordinals in the corpus correspond to 'first', 'second' and 'third'. The following ordinals 'fourth' through 'sixth' were also either recorded or easily elicited. The particularly high occurrence of these forms should not come as a surprise; whereas *pimer* 'first' doubles as a temporal adverb (see 6.3.3 above), the enhanced retrievability of ordinals *sigũd* 'second', *terser* 'third', *kwɔrt* 'fourth', *kĩt* 'fifth' and *sest* 'sixth' probably has to to do with the fact that these forms (slightly transformed in the case of *terser*) feature in the names of weekdays 'monday' through 'friday' (see 8.5.2 for the weekday paradigm). The forms in this small subset are largely suppletive. Some speakers

---

[13]One will notice that, with the exception of 'two hundred' and 'five hundred', all of these are perfectly segmentable into a single numeral plus *sẽt*, which raises the question whether or not these are compound forms or simple attributive constructions. Phonology provides some clues as to the nature of such expressions: whereas auditory stress is inconclusive (by virtue of its being so inconspicuous), the fact that the vowels in *sɛt* and *nɔv* resist rising indicates they must be stressed (see 5.1.1) and therefore supports the interpretation of these forms as the combination of two separate lexemes.

[14]These are standard labels in Indian languages, spelt *lakh* and *crore* in their official latinised forms.

rather hesitantly form further ordinals through the addition of the suffix -*m* to the required numeral[15]: e.g. *oyt* 'eight' > *oytm* 'eighth'. These forms are extremely rare and perceived as acrolectal. English numerals and ordinals have largely replaced their DIP counterparts in the daily usage of DIP speakers.[16]

8.5 Temporal reference

This section is somewhat eclectic, covering purely lexicographic issues as well as the cognitive representation of time and space, and the pragmatics of time reference. Temporal reference being a constant in the communicative instances recorded, the aspects described in this section are essential for proper interpretation of the corpus.

8.5.1 Temporal deixis

In addition to their spatial and discourse deictic role (see 6.5.4), the various demonstratives are also used in temporal phrases. In this context, their semantics reveals a conception of time in which the Present moment is seen as close at hand (and therefore indicated with the proximal demonstrative *es*), the Past is construed as visible but distant (using the distal *ikəl*) and the Future falls outside the temporal deictic space (making use of the absentive *ot(r)*). This is exemplified in (333):

(333) a. *es   an   elz ain nã   vey.*
          DEM$_p$ year 3p still NEG$_{cl}$ come.PST
          'This year they haven't come yet.'

       b. *ikəl   an   tiŋ   muyt chuv.*
          DEM$_d$ year EXS.PST much rain
          'Last year there was a lot of rain.'

       c. *yo a   vi   otr   an.*
          1s IRR.NPST come.INF other year

---

[15] See also section 8.2.3. The suffixes -*imo* and -*ima* are indeed involved in the construction of certain ordinals in Portuguese, but they are highly selective and usually trigger changes in the preceding stem as well (such as in the case of *cem* 'hundred' /*centésimo* 'hundredth').

[16] The preference for English numerals and ordinals instead of earlier systems is to be observed in spoken Gujarati as well, as also in other Indian languages. In Gujarat, it is not uncommon to hear English numerals inserted into an otherwise entirely Gujarati utterance for the indication of prices, age or telephone contacts. In this respect, then, DIP has in the recent past followed a borrowing trend that is regional (or even national) rather than innovative. On the other hand, the spread of English numerals (and lexicon) must be a relatively recent phenomenon among the IP varieties of Diu and Daman, accompanying the crescent interest of the educational system of India in this language; an observer reported, in 1987, that the speakers of Daman IP often replaced Portuguese-derived numerals (as well as lexical labels of family relations, interjections and formulas) with their Gujarati (rather than English) equivalents (Matos 1987:332-335).

'I will come back next year.'

Notice that the primary interpretation of such time expressions (unless overruled by contextual information) is fixed: *ikəl + X* refers to the last occurrence of X, and *ot(r) + Y* refers to the next occurrence of Y. Days have specific labels with reference to the moment of the utterance, viz. *oj* 'today', *õt* 'yesterday' and *amiŋã* 'tomorrow'.

In order to refer to different periods of time, DIP makes use of a construction which could be literally translated as 'after X years/months/...'. In (334a), the moment of reference is the deictic moment of utterance, whereas in (334b) the discourse time is placed in the past:

(334)   a.   *yo ad   vi   voltad   dəpəy də doy trey   an.*
             1s IRR.NPST come.INF return.PTCP after of two three year

             'I'll come back in two or three years.'

        b.   crocodile *dis   use vay,   dəpəy də trey   di   use vẽy*
             crocodile say.PST 2   go.NPST after of three day 2   come.NPST
             volta-d.
             return-PTCP

             'The crocodile said: "Go and come back in three days".'

8.5.2   Days

The DIP word for 'day' is *di* [from Ptg. *dia* 'day'], and *sɛman* [from Ptg. *semana* 'week'] refers to a week. The names of weekdays in DIP are as follows:

- *dumĩg̃* 'sunday' [from Ptg. *domingo* 'sunday']
- *sigũd fer* 'monday' [from Ptg. *segunda-feira* 'monday']
- *ters fer* 'tuesday' [from Ptg. *terça-feira* 'tuesday']
- *kwɔrt fer* 'wednesday' [from Ptg. *quarta-feira* 'wednesday']
- *kĩt fer* 'thursday' [from Ptg. *quinta-feira* 'thursday']
- *sest fer* 'friday' [from Ptg. *sexta-feira* 'friday']
- *sab(d)* 'saturday' [from Ptg. *sábado* 'saturday']

With respect to the weekday paradigm, DIP follows very closely the Portuguese model, which uses compound labels for all days other than 'saturday' and 'sunday' combining an ordinal and the noun *feira* 'fair/market'. For the role of ordinals in the DIP weekdays, see 8.4.2. English weekday names substitute for this Portuguese-derived paradigm very often in actual speech.

8.5.3   Months and seasons

The DIP word for 'month' is *mes* [from Ptg. *mês* 'month'], and that for 'year' is *an* [from Ptg. *ano* 'year']. Like weekdays, speakers of DIP tend to use a Portuguese-derived set of labels for the various months, although their English equivalents are very common. The Portuguese-based paradigm is given below:

- *janer* 'january' [from Ptg. *janeiro* 'january']
- *fɛvrer/fɛvirer* 'february' [from Ptg. *fevereiro* 'february']
- *mars* 'march' [from Ptg. *março* 'march']
- *abril* 'april' [from Ptg. *abril* 'april']
- *may* 'may' [from Ptg. *maio* 'may']
- *juŋ* 'june' [from Ptg. *junho* 'june']
- *juy* 'july' [from Ptg. *julho* 'july']
- *agos* 'august' [from Ptg. *agosto* 'august']
- *sɛtem(b)* 'september' [from Ptg. *setembro* 'september']
- *otub* 'october' [from Ptg. *outubro* 'october']
- *nɔvem(b)* 'november' [from Ptg. *novembro* 'november']
- *dɛzem(b)* 'december' [from Ptg. *dezembro* 'december']

The three local seasons of the year are referred to as follows:

- *chuv* [Lit. 'rain'] or *ĩver* 'rainy season/monsoon' [from Ptg. *chuva* 'rain' and *inverno* 'winter']
- *fri* [Lit. 'cold'] or *tẽp də fri* [Lit. 'time of cold'] 'winter' [from Ptg. *(tempo do) frio* 'cold (season)']
- *kalor* [Lit. 'heat'] 'summer' [from Ptg. *calor* 'heat']

8.5.4   Telling time

The DIP questions used to elicit information of clock time are given in (335); of the two formulations, (335a) is the most common:

(335)   a.   *kwõt   ɔr   (jə)   fik-oʔ*
             how.much hour already become-PST
             'What time is it?' [Lit. 'How many hours have become?']

        a.   *kwõt   ɔr   ɛʔ*
             how.much hour COP$_i$.NPST
             'What time is it?'

The reply is usually rather straightforward, consisting of the coordination of the figure for the hours (optionally followed by *ɔr* 'hour') and the figure for the minutes (optionally followed by *minut* 'minute'). Therefore, for instance, *kwɔt i doy* 'four hours and two

minutes', or *sĩk ɔr i vĩt i doy* 'five hours and twenty-two minutes'. Special labels include *midi* 'midday' and *menoyt* 'midnight' (both equivalent to *doz ɔr* 'twelve hours'), and also *me/mey* 'half' for a thirty minute figure (e.g. *dɛz i me* 'half past ten').

Some numerals, behave somewhat differently in such contexts, suggesting a certain degree of crystallisation. This refers particularly to those numerals whose etymons contain a final sibilant segment, which in DIP is usually not realised except in these constructions. Therefore, *duz i me* and not *doy i me 'half past two', also *duz ɔr* and not *doy ɔr 'two o'clock'[17], also *trez ɔr* and not *trey ɔr 'three o'clock'.

8.6 Semantic fields

This section explores the characteristics of four semantic fields, self-contained case studies which reveal various important etymological links, semantic patterns, and processes shaping the lexicon of DIP. The semantic fields under analysis are lexemes denoting human categories, including family relations (section 8.6.1), colours (in section 8.6.2), animals (section 8.6.3) and toponyms (in section 8.6.4).

8.6.1   Human referents

The lexicon referring to human categories and human relations is the only one that consistently (though not universally) features lexicalised gender distinctions in DIP, reflecting natural gender. The pairs below are very common in DIP:

- *avo* 'grandfather' vs. *avziŋ* 'grandmother' [from Ptg. *avô* 'grandfather' and *av-* + *-inha* 'grandparent + DIM respectively; see 8.2.1']
- *baba* 'infant boy' vs. *bai*[18] 'infant girl' [attributed by Dalgado (1900) to common Indo-Aryan terms][19]
- *irmãw* 'brother' vs. *irmã* 'sister' [from Ptg. *irmão* 'brother' and *irmã* 'sister' respectively]
- *marid* 'husband' vs. *muyɛr* 'wife' [from Ptg. *marido* 'husband' and *mulher* 'woman' respectively]
- *nɛt* 'grandson' vs. *nitiŋ* 'granddaughter' [from Ptg. *neto* 'grandson' and *net-* + *-inha* 'grandchild + DIM' respectively; see 8.2.1]
- *ɔm* 'man' vs. *muyɛr* 'woman' [from Ptg. *homem* 'man' and *mulher* 'woman' respectively]

---

[17]In the case of *duz*, the explanation for the difference observed in clock time expressions involves the fact that, in Portuguese, the cognate numeral inflects for gender: Ptg. *dois* 'two [m.]', *duas* 'two [f.]'. In similar time expressions, it is the feminine form to occur given that *hora* 'hour' is feminine. It seems, therefore, that DIP has crystallised this form in clock time expressions, given that this form of the numeral never occurs elsewhere.

[18]Instances of both *baba* (or *babá*) and *bai* (or *baî*) are recorded in Schuchardt (1883) and Quadros (1907:192*ff*).

[19]Widely used in Konkani, as reported by Furtad (see 1999); see also 8.7.2.

- *papa* 'older family relation' [m, respectful] vs. *mama* 'older family relation' [f, respectful] [unclear origin]
- *pay* 'father' vs. *mãy* 'mother' [from Ptg. *pai* 'father' and *mãe* 'mother' respectively]
- *rapaz* 'boy' vs. *rap(ə)rig* 'girl' [from Ptg. *rapaz* 'boy' and *rapariga* 'girl' respectively]
- *siŋor* 'Mr.' vs. *don* 'Mrs.' [from Ptg. *senhor* 'Mr.' and *dona* 'Mrs.' respectively]
- *uncle* 'older man' vs. *auntie* 'older woman' [from Eng.]

Other frequent members of the subset do not formalise a gender distinction or enter a gender opposition. These include:

- *kuŋad* 'brother/sister-in-law' [from Ptg. *cunhado* 'brother-in-law' or *cunhada* 'daughter-in-law']
- *fil* 'son/daughter' [from Ptg. *filho* 'son' or *filha* 'daughter']
- *noiv* 'fiancé/e' [from Ptg. *noivo* 'fiancé' or *noiva* 'fiancée']
- *prim* 'cousin' [from Ptg. *primo* 'cousin [m.]' or *prima* 'cousin [f.]']
- *subriŋ* 'nephew/niece' [from Ptg. *sobrinho* 'nephew' or *sobrinha* 'niece']
- *tɛtɛ* 'auntie' [unclear origin]
- *ti* 'uncle/aunt' [from Ptg. *tio* 'uncle' or *tia* 'aunt']

8.6.2   Colours

Colours are an interesting subset of the lexicon in that, even though most words are derived from Portuguese etyma, it features some innovations based on the local context. The following are clearly derived from Portuguese lexemes:

- *amrɛl* 'yellow' [from Ptg. *amarelo* 'yellow']
- *azul* 'blue' [from Ptg. *azul* 'blue']
- *brãk* 'white' [from Ptg. *branco* 'white']
- *larãj* 'orange' [from Ptg. *laranja* 'orange']
- *pret* 'black' [from Ptg. *preto* 'black']
- *(kod)roz* 'pink' [from Ptg. *(cor de) rosa* 'pink']
- *verd* 'green' [from Ptg. *verde* 'green']
- *vɛrmey* 'red' [from Ptg. *vermelho* 'red']

The following colours depart from the Portuguese model:

- *kafɛ* 'brown' [from DIP *kafɛ* 'coffee']
- *jamlãw* 'purple' [from DIP *jamlãw* 'jamun']

The semantic extension *kafɛ* and *jamlãw* is modelled on Gujarati, which uses the equivalent of 'coffee' to refer to (dark) brown colour and *jambli* 'jamun' for 'purple'. We are dealing with calques rather than borrowings, given that the DIP forms differ from their Gujarati counterparts, even though they show certain similarities: *kafɛ* is derived from

Ptg. *café* 'coffee', and the *-ãw* ending in *jamlãw* also reveals a Portuguese source.[20]

8.6.3 Animals

This section provides only a small sample of the vast lexicon of DIP linked with animal referents. It is my intention to highlight that this subset of the lexicon (as well as other semantic fields such as food, plants or household objects) features a number of different sources. To introduce this semantic field, I will provide some generic (hypernymic) lexemes:

- *animal* 'animal' [from Ptg. *animal* 'animal']
- *bix* 'bug' [from Ptg. *bicho* 'animal']
- *pas(r)* 'bird' [from Ptg. *pássaro* 'bird']
- *pex* 'fish' [from Ptg. *peixe* 'fish']

It is not uncommon for competing forms from different sources to be used frequently. We therefore find that Portuguese-derived *liãw* 'lion' and English *lion* are both well represented in the corpus, as are Portuguese-derived *pavãw* 'peacock' and English *peacock*. All in all, however, Portuguese-derived terms are still dominant. One interesting subset of this semantic group consists of items derived from Portguese etyma which were common in Classical Portuguese but have disappeared from the active lexicon of modern SP (see also 8.7.1). Those include the nouns below:

- *adiv* 'fox, jackal' [from Classical Ptg. *adive* or *adibe* 'jackal']
- *mãduk* 'frog' [from Classical Ptg. *manduco* 'frog']

Other nouns, such as the following, are derived from Portuguese etyma still in present-day SP:

- *bur* 'donkey' [from Ptg. *burro* 'donkey']
- *gray* 'crow' [from Ptg. *gralha* '[a type of] crow']
- *karẽgej* 'crab' [from Ptg. *caranguejo* 'crab']
- *kasãw* 'sharkfish' [from Ptg. *cação* '[small] shark']
- *muskit* 'mosquito' [from Ptg. *mosquito* 'mosquito']
- *pavãw* 'peacock' [from Ptg. *pavão* 'peacock']
- *pork* 'pig' [from Ptg. *porco* 'pig']

A crucial observation is that part of the DIP lexicon is not derived from Gujarati etyma but from Konkani (the Indo-Aryan language of Goa and some pockets on the Kanara Coast) lexemes - the significance of this fact is explored in 8.7.2. Examples include:

---

[20]The jamun fruit, endemic to India and transplanted to Mozambique and Brazil, is known in Portuguese as *jambolão* or *jamelão*. They are called *jambolões* in Orta (1563). With respect to the crops of Goa, Daman and Diu, Bragança Pereira (1940:150) identifies the *jamboleiro*, described as a 'árvore de fruto, de cujos frutos se faz vinho' [a fruit tree from whose fruits one can make wine].

- *bumli* 'Bombay Duck' [from Konk. *bombil* or *bomble* 'Bombay Duck']
- *chani* 'squirrel' [from Konk. *chani(m)* 'squirrel']
- *chichŭdri* 'mouse' [from Konk. *chichandor* 'mouse']
- *gãdil* 'wasp/bee' [from Konk. *ganzil* 'honeybee']

Finally, some animal names are clear English loans, and not always phonologically adapted to what is determined in chapter 5 to be the core phonology of DIP lexemes:

- *crocodile*
- *crow*
- *lion*
- *tortoise*

8.6.4 Toponyms

For proper interpretation of this section, refer to the maps in chapters 1 and 3. Toponyms in DIP have not fully integrated with the official names of places in Gujarati, therefore retaining to a large extent previous appelations from the period of Portuguese colonial rule. For instance, while the village the Portuguese called *Brancavará* is now officially known as *Vanakbara*, DIP retains the form *Brãkavara*. Similarly, the Diuese enclave known today as *Simar* is known to the speakers of DIP as *Sĩbor* (*Simbor* for the Portuguese). The village of *Goghla* is known in DIP as *Gogla*, with no aspiration of the plosive segment. On the other hand, the erstwhile Portuguese toponym *Podamo* has not survived in modern DIP, as all speakers refer to the village as *Fudam* after its new official name. Other toponyms on the island include *Buxvara* (as opposed to the modern name *Bucherwara*), for which the Portuguese used *Buxinvará*.

Diu Town is known simply as *Diw* in DIP, although it can be distinguished from the remainder of the island with the epithet *pras* [from Ptg. *praça*, used in Classical Portuguese to refer to a colonial stronghold]. The urban area surrounding the churches, nowadays known in general as *Firangiwara* [from Guj. *firangi* 'foreigner' and *wada*, a common element in toponyms] can also be known among the speakers of DIP as *bar də kristãw* 'neighbourhood of the Christians'. In DIP, the coastal stretch in the city centre is variably referred to as *bãdər* [from Guj. *bundər* 'port'] or *pray* 'beach' (even though no beach presently exists). Outside the city, the two spots which allow for safe crossing onto the mainland retain their Portuguese names in DIP: *pas sek* [from Classical Portuguese *passo* 'ford' and *sêco* 'dry'] and *pas kov* [from Ptg. *passo* and *côvo* 'deep'].[21]

Within the city, the main fort is still known as *kastɛl* 'castle' [the official nomenclature in Classical Portuguese] and the sea fortress (known in Gujarati as *Panikota* 'water fort')

---

[21] As recorded in Bragança Pereira (1940:22), in his description of the territory of Diu: 'banhada ao Sul pelo Gôlfo da Arábia e separada do continente por um braço de mar que constitui o pôrto com as suas duas barras, uma em frente da fortaleza e a outra em Brancavará. O esteiro, nas Marés baixas dá vau em dois pontos: Passo Sêco e Passo Côvo.' [bathed on the south by the Arabian Gulf and separated from the continent by an inlet constituting the port with its two anchorages, one in front of the fortress and the other in Brancavará. During low Tide, the river can be vaulted at two spots: Passo Sêco and Passo Côvo].

is called *furtiŋ dɛ mar* or *fɔrt dɛ mar*. The city gates also retain their Portuguese names: *Pɔrt də kăp* 'Door of the fields' and *Pɔrt də fɔr* 'Outer Door' (see Map 3.2 in section 3.1).

In the immediate vicinity of the territory, the closest towns are known to the speakers of DIP by their Gujarati names: e.g. *Una*, *Veravəl*, *Kohinar* (see chapter 3, the map in Figure 3.1). Further afield, four toponyms of particular relevance for the Catholics of Diu retain forms derived from their Portuguese counterparts: *Damãw*[22] 'Daman' [Ptg. *Damão*], *Boməĩ* 'Mumbai/Bombay' [Ptg. *Bombaim*], *Go* 'Goa' [Ptg. *Goa*] and *Silvas* 'Silvassa'[23] [Ptg. *Silvassa*]. Several locations in Goa are also referred to with Portuguese-derived names, as they are in Goan Portuguese: e.g. *Pãjĩ* 'Panaji' [Ptg. *Pangim*], *Margãw* 'Madgaon' [Ptg. *Margão*]. The most common toponyms referring to the territory of Portugal are *purtəgal* 'Portugal' and *lisbo* 'Lisbon' [Ptg. *Lisboa*].

8.7 Notes on etymology

This section does not intend to ascertain the full etymological ancestry of the DIP lexicon, but simply to expand on two interesting points revealed in the previous study of particular semantic fields. The first of these is the fact that DIP preserves various words derived from etyma which have become archaisms in modern European Portuguese; this is dealt with in section 8.7.1. Section 8.7.2 accounts for the observed Konkani element, while sections 8.7.3 and 8.7.4 explore various means by which Gujarati helped shape the lexicon of DIP.

8.7.1 Portuguese archaisms

Apart from the words *adiv* 'fox, jackal' and *mãduk* given in section 8.6.3 (from Classical Ptg. *adive/adibe* 'jackal' and *manduco* 'frog' respectively), the recorded lexicon of DIP contains various other words which were current in Classical Portuguese but became archaisms in Modern Portuguese. These include:

- *alpak* 'sandal, slipper' [from Classical Ptg. *alparca* 'sandal']
- *arəma* 'to hit' [from Classical Ptg. *arrimar* 'to hit, to beat, to lean against']
- *chumas* 'pillow' [from Classical Ptg. *chumaço* 'pillow']
- *kavok* 'cave, hole' [from Classical Ptg. *cavouco* 'cave']
- *patrater* 'pompous' [from Classical Ptg. *patarateiro* 'pompous']
- *mufin* 'avaricious, sad' [from Classical Ptg. *mufino* 'avaricious, sad']
- *sŭbrer* 'umbrella' [from Classical Ptg. *sombreiro* 'umbrella']

Another conspicuous semantic distinction between Classical Portuguese (as well as dialectal Portuguese in rural environments) and modern standard Portuguese concerns the

---

[22] In DIP, as in Daman IP, the two fortified areas of Daman town are known as *Damãw Gran* 'Big Daman' [Ptg. *Damão Grande*] and *Damãw Piken* [Ptg. *Damão Pequeno*] 'Small Daman', where Gujarati uses the semantically equivalent *Moti Daman* and *Nani Daman* respectively.

[23] The main town in Dadra and Nagar-Haveli, the two enclaves now integrating the Union Territory of Daman, Diu, Dadra and Nagar-Haveli - see chapter 1.

names of the daily meals. In Classical Portuguese *almoço* was the morning meal ('break-fast'), *jantar* referred to the midday meal ('lunch') and *ceia* was the evening meal ('dinner'). In time, with the change of eating habits, the term *pequeno-almoço* was introduced to sig-nify 'breakfast', *almoço* became 'lunch', *jantar* became 'dinner' and *ceia* was reserved for irregular, late evening meals. The DIP words derived from these etyma retain a semantic value closer to the archaic Portuguese system: DIP *almos* 'breakfast', *jãtar* 'lunch' and *sey* 'dinner'.

The retention of Portuguese archaisms in DIP is highly significant but not unexpected. It indicates that, for all the change DIP must have undergone throughout its history, at least part of its lexicon was established early on and was not replaced with the modern alternatives of SP. See section 9.7.2 for a discussion of the theoretical significance of the fact.

8.7.2   The Konkani element

As shown in previous sections, there is a significant Konkani element which permeates the DIP lexicon. Apart from the human categories and animal nouns described in 8.6.1 and 8.6.3 respectively, there are several other points of contact between the two languages. For instance, the DIP expression *tok bok* [Lit. 'touch mouth'], which refers to a 'small snack' or a 'meal' is also used in Goan Portuguese (*toca-bocca*) and it is traced back by Dalgado (1900) to the Konkani translational equivalent *tondák lāvunchem*.

This fact is highly significant; given that all native speakers of DIP are fluent speakers of Gujarati, and only a minority have any significant knowledge of Konkani, it seems odd that they should rely on Konkani-derived (rather than Gujarati-derived) lexemes to this extent. On the other hand, one must be aware of the long-standing cultural, social and even political connection between Diu and Goa (see chapter 3).

There are two possible scenarios to account for the Konkani element in DIP, which I will succintly present below:

- The first scenario posits that Konkani-derived lexemes may have been part of the DIP lexicon from its very inception, given that Goa became, even before the colonisation of Diu, the headquarters of the Portuguese efforts in the region and the great population platform of the *Estado da Índia* (providing slaves, officials, military, missionaries and set-tlers to the remaining Portuguese possessions). There is a body of lexemes found across the various discrete Indo-Portuguese varieties, and even other Portuguese-lexified creoles of Asia,[24] constituting a sort of *common Asian-Portuguese vocabulary*, to which Konkani is likely to have contributed significantly. Unfortunately, the application of this hypothesis to Diu cannot be verified in the absence of early records of DIP.

- The second explanation ties in with the fact that, until very recently, SP - albeit with some peculiarities including certain Konkani-derived lexemes (see Dalgado 1900) - was rel-

---

[24]Some examples include *ap/apa* 'flatbread' originally from Tamil *appam* (Dalgado 1913); *bibik/bebinca* '[a layered cake]', of unknown origin; *jag/jagra* 'jaggery', originally from Konk. *sākar* or Sanskrit *xarkarā* (Dalgado 1900); *jak/jaca* 'jackfruit', originally from Mallayalam. *chakka* (Dalgado 1913); *mat/mate* 'earth-enware, clay', originally from Prakrit *māti* or Sanskrit *mrittiká* (Dalgado 1900); *maynat/maynato* 'clothes' washer', originally from Mallayalam *maṇṇatán* (Dalgado 1913).

atively widely used in Goa. It is therefore possible that Konkani lexemes were transferred onto DIP through Goan Portuguese (the language spoken by many officials and clergy in Diu in the recent past) rather than directly from Konkani. This suggests that in fact the modern *norm* (the privileged source of lexemes among native speakers of DIP) may well be the Portuguese of Goa and not necessarily that of Europe.

Any of these explanations would be sufficient on its own, but in truth they are not mutually exclusive. Historical data (see section 3.3) does confirm the realisation that Goa, and Konkani speakers, have continuously played an important role in Diuese society, in the 16th century as much as the 21st (see 3.3).

8.7.3   The lexical contribution of Gujarati

Gujarati, unsurprisingly, is an important source of lexical material for the core vocabulary of DIP. Apart from various lexemes mentioned earlier (e.g. *avŭ* 'so, thus', *bãdər* 'port' *giligili* 'tickle', *chipia/chipio* 'tweezers', *kanjus* 'avaricious, miser'), various other common DIP words are borrowed from Gujarati, in particular those referring to local entities and concepts. Some examples recorded in the corpus include:

- *bajəri* 'millet' [from Guj. *bajri* 'millet']
- *boro* 'sweet dish with coconut' [from Guj. *bolo* 'sweet dish']
- *bud* 'ghost' [from Guj. *budh* 'ghost']
- *chalu* 'functioning, on' [from Guj. *chalu* 'functioning, on']
- *dal* 'lentils' [from Guj. *dal* 'lentils']
- *mãdir* 'Hindu temple' [from Guj. *mandir* 'Hindu temple']
- *majid* 'mosque' [from Guj. *masjid* 'mosque']

Certain functional items have also been borrowed into DIP, often carrying over their original functional load. One case in point is the DIP requestative particle *nə*, which is entirely equivalent, in formal as well as functional terms, to the Gujarati element *nə*. The functions of this particle in Gujarati are described as follows:

- in Imperatives: in Gujarati, the apposition of *nə* to an imperative verb results in a more respectful and yet more insistent request - e.g. *bes!* 'sit down' vs. *bes nə!* 'do sit down';

- in Interrogatives: added to clause which is formally equal to an affirmative, *nə* results in a polar question (i.e. yes/no question) which expects confirmation and may be entirely rhetorical;

- Pragmatic marker: *nə* is often used to ascertain the identifiability of an element introduced to discourse, in a construction of the type: X *nə*, DEM$_X$ V.

A comparison between the various uses of *nə* in Gujarati and DIP (described in sections

7.6.2.2 and 7.6.3.1) reveals a complete overlap. This leaves no doubt that Gujarati is the source of the requestative particle in DIP, even though one might entertain the possibility that it had derived from the common Portuguese tag *não é > né* 'isn't it'.

8.7.4 The semantic influence of Gujarati

In the previous section, we have seen instances of direct lexical contribution from Gujarati, but there is another mechanism through which Gujarati has left a strong imprint in the lexicon of DIP. In many cases, the semantic scope of elements derived from different sources (chiefly from Portuguese) was extended or reduced to fit Gujarati patterns.

We have already discussed, in 8.1.3, the cases of DIP *pε* and *mãw*, which refer to 'foot/leg' and 'hand/arm' respectively in tune with Gujarati. We also discussed the semantic scope of the DIP verb *fika*, which ranges from 'to stay' to 'to dwell' and 'to become'. The range of meaning of this verb owes as much to Portuguese - in which *ficar* has the meanings 'to stay' and 'to become' - as to Gujarati - in which *rahevũ* means both 'to stay' and 'to dwell' (Masica 1991:391).

A striking example of the Gujarati semantic influence refers to the partial reinterpretation of the (former) DIP diminutive suffix -*iŋ* as a gender marker (see 8.2.1). In Gujarati, the three genders may code semantic differences of size and coarseness when applied to inanimates. Masculine-inflected nouns are expected to refer to large objects while feminine-inflected nouns refer to smaller entities - effectively resulting in a diminutive construction. Neuter nouns can refer to unusually large and/or coarse objects, normally with a pejorative sense. Consider the following examples, given in Cardona (1965), Cardona and Suthar (2003) and graphed here accordingly:[25]

- *roṭlo* 'thick bread' vs. *roṭli* 'roṭī (bread)' vs. *roṭlũ* 'coarse bread (pejorative)'
- *cəmco* 'large spoon' vs. *cəmci* 'teaspoon'
- *gəḷni* 'tea strainer' vs. *gəḷnũ* 'filtering cloth'
- *nɔkri* 'job, service' vs. *nɔkrũ* 'job that is looked down upon'

The use of feminine suffixes with diminutive semantics is observed across the Indo-Aryan language family (Masica 1991:77-78). This superimposition of feminine and diminutive semantics helps to explain the general reinterpretation of the Portuguese diminutive suffix -*inho/inha* observed in DIP as well as other Indo-Portuguese varieties.[26] In DIP, -*iŋ* marks the feminine-gender element in such natural gender pairs as *nεt* 'grandson' vs. *nitiŋ* 'granddaughter', *avo* 'grandfather' vs. *avziŋ* 'grandmother'.

---

[25] In Gujarati, the typical masculine gender suffix is -*o*, while feminine is marked with -*i* and neuter with -*ũ*. See 9.2.1 for further information.

[26] In his study of the *Norteiro* (Bombay area) variety of Indo-Portuguese, Dalgado (1906) also reported the feminine reading of certain seemingly diminutive forms, such as *noivinho* 'bride'.

Part III

Discussion

Chapter 9

Comparative study

The previous chapter has already established that the overwhelming majority of DIP words (be they lexical or purely functional) is derived from Portuguese. It is now necessary to take the analysis beyond this stage and explore to what extent DIP aligns with either Gujarati or Portuguese in typological terms, and to what extent it diverges from both. This chapter therefore carries out a comparative study focusing on selected aspects of phonology, phonetics and morphosyntax, and contrasting DIP with both Gujarati and (Classical) Portuguese. The features explored are intended as illustrations only, given that a full comparative study is beyond the scope of this work. It is therefore important to clarify that the evidence provided here cannot be used to quantify the relative contribution of Gujarati and Portuguese to the structural make-up of DIP.

There is also a diachronic side to this chapter attempting to establish, to the extent possible, the origin of particular DIP features and its developmental chronology. DIP has been in sustained contact with both Gujarati and Portuguese for over four centuries, which makes it a privileged testing ground for the hypothesis that the initial moments of contact are particularly decisive in establishing the features of a high-contact variety - see section 9.7 for further discussion. Linguistic change being all-pervasive, affecting all the languages involved in the equation, it is relevant to establish whether the drift of DIP has been tied onto change affecting Portuguese or Gujarati.

The first section (section 9.1) contrasts DIP phonology and phonetics - as described in chapter 5 - with the Portuguese and Gujarati systems, from a diachronic perspective. Section 9.2, on morphology, deals specifically with nominal and verbal inflection. A comparative study of personal pronoun, possessive and deictic paradigms can be found in 9.3, while 9.4 is dedicated to the issue of basic word order. Section 9.5 is rather specific, contrasting the three languages with respect to certain aspects of case-marking. A selection of syntactic constructions and precepts is analysed in 9.6, including the syntax of content questions, copular and comparative constructions, passivisation, complex predicates, negative concord and the form of utterance complements. Finally, 9.7 draws some conclusions from the comparative exercise.

9.1 Phonology and phonetics

In order to assess the contribution of any language to the formation of a contact variety, we must know not only the present states of the relevant languages but also their past stages; it would be convenient to have an idea of the exact sort of input available in early 16th-century Diu, but our knowledge of the past stages of languages is usually bound to written sources and therefore to relatively prestigious varieties which may not necessarily correspond to the input for the formative pool of high-contact languages. Gujarati diachronic phonology is not as developed as desirable; for European Portuguese, on the other hand, some information is available not only on the (more or less) standard varieties of the 16th century but also on dialectal variation. We will see that the present-day phonology of DIP retains certain features from Portuguese that are nowadays either lost or retracted to dialectal speech; in this respect the study of DIP contributes to advance our knowledge of the chronology of phonological change in Portuguese.

This section is organised as follows. Subsections 9.1.1 and 9.1.2 will in turn describe the phonologies of Portuguese and Gujarati, both past (to define the initial input) and present (to assess their patterns of change) while establishing links with synchronic DIP phonology. 9.1.3 draws some conclusions concerning the influence of both Gujarati and Portuguese onto the phonology of DIP.

9.1.1 Diachronic phonology of Portuguese

The vocalic system of Classical Portuguese does not differ greatly from Modern Portuguese, except with respect to the unstressed vowels. Figure 9.1 charts the vocalic system of Classical Portuguese. All segments were allowed in atonic position, and all except the one enclosed in brackets occurred in stressed position.[1]
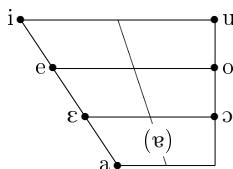

Figure 9.1: Oral vocalic inventory of Classical Portuguese (adapted from Castro 1991)

The nasal vowel inventory of Classical Portuguese[2] was reduced in relation to the array

---

[1] 16th-century grammarians such as Fernão de Oliveira (1536) and Duarte Nunes de Leão (1576) already noted that, while [ɐ] did occur in stressed position, it was an allophone of /a/ licensed by a subsequent nasal phoneme.

[2] Written sources cannot be conclusive as to the actual pronunciation of all these segments in Classical Portuguese; if, for instance, the low nasal phoneme /ã/ were realised as [ẽ], the system would be equivalent to that of Modern Portuguese - see Figure 9.4 below. It is more likely that the phoneme had a low realisation, however, as this segment occurs in a number of Portuguese-lexified creoles and it is retained in

of oral vowels, charted in Figure 9.2:
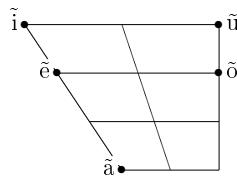
Figure 9.2: Nasal vocalic inventory of Classical Portuguese (adapted from Teyssier 2001)

As far as the vocalic segments go, the most conspicuous change in Portuguese since the 16th century has to do with the unstressed vowels. Across the centuries, the unstressed mid back vowels /ɔ/ and /o/ have risen to /u/, the unstressed low vowel /a/ has risen to /ɐ/, and the unstressed mid front vowels /ɛ/ and /e/ have centralised to /ə/. Very few exceptions to this tendency remain in Modern European Portuguese. The same considerations on the occurrence of [ɐ] in stressed position still apply, along with some additional contexts such as in a diphthong /ɐj/. Figure 9.3 charts the oral vocalic system of Modern Portuguese. All vowels *outside* brackets occur in tonic position, and all *italicised* segments are allowed in unstressed position.
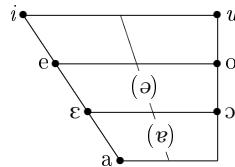
Figure 9.3: Oral vocalic inventory of Modern Portuguese

The nasal vowels of Modern Portuguese are given in Figure 9.4.
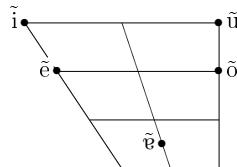
Figure 9.4: Nasal vocalic inventory of Modern Portuguese

---

some of the most conservative dialects of northern Portugal (see Mateus and d'Andrade 2000).

The only difference of the modern-day inventory when compared to that in Figure 9.2 is that, in the standard variety, /ã/ has been replaced with /ẽ/. Northern Portuguese dialects still retain the previous low realisation of the phoneme.

Table 9.1 below presents the phonemic system of Portuguese around 1500, as reconstructed from written sources.

	Bilabial	Lab-dent.	Dental	Alveolar	P-alveo.	Retroflex	Palatal	Velar	Uvular
Plosive	p b		t d					k g	
Nasal	m			n			(ɲ)		
Trill				r					
Tap/Flap				ɾ					
Fricative		f v		s z	ʃ				
				s̺ z̺					
Affricate				t͡ʃ d͡ʒ					
Approx							j		
Lab. appr								w	
Lateral				l				ʎ	

Table 9.1: Consonant inventory of Classical Portuguese, (adapted from Teyssier 2001)

It would be imprudent to assume old written sources could ever capture the full dialectal range of Portuguese around 1500; furthermore, it is not always evident when a particular change began and when it was completed. One such process of change, as will be seen in section 9.1.3, relates to the exact period in which the affricate /d͡ʒ/ gave place to the fricative /ʒ/. Data from contact varieties such as DIP, for which a clear chronological starting point can be advanced, turn out to be essential to clarify such questions; in this particular case, I have decided to include the affricate segment in Table 9.1 instead of the fricative precisely because modern DIP data shows the affricate is the only relevant etymological variant.

One further comment concerning the system of the fricatives is that, by the time contact between Gujarati and Portuguese began in Diu, a widespread wave of simplification had already cancelled the plosive release of Old Portuguese [t͡s] and [d͡z] and in the process had given rise to an opposition between two close sets of fricatives: /s/ and /z/, on the one hand, and apico-alveolar /s̺/ and /z̺/ on the other (see Cunha and Cintra 1984).

In early Classical Portuguese, /ɲ/ had a much more limited distribution than it does in Modern Portuguese, the reason for which it is bracketed in Table 9.1. This segment was epenthetically added to resolve several hiatuses involving a nasal vowel other than /ũ/ and an oral vowel. A full discussion of the phenomenon and its significance for the formation of DIP can be found in section 9.1.3.

The alveolar lateral segment of Classical Portuguese velarised in syllable-final position (cf. Teyssier 2001:102), which is still the case at present.

Table 9.2 below presents the inventory of Modern Portuguese consonants with phonological status.

	Bilabial	Lab-dent.	Dental	Alveolar	P-alveo.	Retrof.	Palatal	Velar	Uvular
Plosive	p b			t d				k g	
Nasal	m			n			ɲ		
Trill				r					ʀ
Tap/Flap				[ɾ]					
Fricative		f v		s z	ʃ ʒ				
Affricate									
Approx.							j		
Lab. appr.								w	
Lateral				l			ʎ		

Table 9.2: Consonant inventory of Modern Portuguese

A comparison of Tables 9.1 and 9.2 reveals that the fricative system of Portuguese has simplified somewhat. The four way alveolar paradigm has reduced to two segments[3], resulting in /s/ and /z/ for standard Portuguese and /s̠/ and /z̠/ for the central dialects. Affricates are now also exclusive to dialectal speech; in standard Portuguese, the transformation of Old Portuguese /ʤ/ into /ʒ/ is complete, while Old Portuguese /ʧ/ and /ʃ/ have fused.

A uvular articulation of the trill has become widespread in standard Portuguese in optional alternation with its alveolar counterpart. /ɲ/ has gained more prominence in Modern Portuguese.

9.1.2 Diachronic phonology of Gujarati

The diachronic exercise done for Portuguese in the previous section can unfortunately not be completely replicated for Gujarati, as there is little accessible information concerning Middle (i.e. 15th-century through 18th-century) Gujarati. The data presented in Figure 9.5 therefore concerns modern standard Gujarati only.
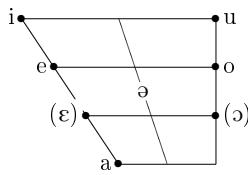

Figure 9.5: Oral vocalic inventory of Modern Gujarati

The brackets in Figure 9.5 are meant to show that the mid-high and mid-low vowels only contrast in very limited contexts[4], and that it is /ɛ/ and /ɔ/ which occur more rarely.

---

[3] The dialects of Northeastern Portugal are exceptional in that they still retain the opposition between all four segments.

[4] In certain dialects, no opposition whatsoever obtains between these two sets of vowels (cf. Modi 1994), and the status of the mid-low vowels have long been a matter of debate (see Turner 1925, Pandit 1955)

Vowel length is not contrastive, but differences in length do obtain according to phonetic contexts (e.g. vowels tend to be longer in word-final position and in monosyllables).

Nasality is phonological in Gujarati, and so is breathy voice (*murmur* in traditional descriptions of Gujarati), as minimal pairs do obtain between breathy voiced and homorganic oral vowels. A breathy voiced realisation at the phonetic level can also be caused by the presence of a subsequent aspirated consonant (e.g. the /a/ in *Gandhi*). Saurashtrian Gujarati is notable in that breathy voice is less prevalent than in the Standard.[5] On the other hand, Saurashtrian is said to be the dialect where nasalisation is the most ubiquitous - Standard Gujarati has entered a phase of progressive vowel desanalisation, which does not affect Saurashtrian to the same extent.

Gujarati allows both rising and dropping diphthongs, contrast for instance /jad/ 'memory' with /pəjsa/ 'money' (*apud* Mistry 1997:659). In addition, a final /-i/ or /-u/ after a vowel may be diphthongised, giving rise to pairs such as [rai]/[ray] 'mustard' and [beu]/[beʊ] 'both' (Cardona and Suthar 2003:666).

Table 9.3 presents the phonological inventory of Modern Gujarati, as retrieved from several sources.

	Bilabial	Lab-dent.	Dental	Alveolar	P-alveo.	Retrof.	Palatal	Velar	Glottal
Plosive (aspirated)	p b pʰ bʰ		t d tʰ dʰ			ʈ ɖ ʈʰ ɖʰ		k g kʰ gʰ	
Nasal	m			n		ɳ			
Trill									
Tap/Flap				ɾ					
Fricative				s z		ʂ			h
Affricate (aspirated)				tʃ dʒ tʃʰ dʒʰ					
Approx.		ʋ					j		
Lateral				l		ɭ			

Table 9.3: Consonant inventory of Modern Gujarati

Cardona and Suthar (2003) indicate the postalveolar position as the place of articulation for the flap, which is for them equivalent to retroflex (this should be noted [ɽ] in IPA). Mistry (1997), on the other hand, is very clear in classifying this as an alveolar segment, opposed to the dental plosives. I have therefore opted for his interpretation.

In addition to the three nasal phonemes indicated in Table 9.3, traditional descriptions of Gujarati phonology often propose a fourth nasal phoneme /N/ to account for occurrences of [ɲ] and [ŋ]. However, these appear to be no more than contextual allophones of /n/: [ŋ] occurs between a schwa and a velar plosive (aspirated as well as unaspirated), while [ɲ] is common before fricatives. For this reason, no /N/ is postulated here.

As far as the fricatives are concerned, Mistry (1997) and Cardona and Suthar (2003)

---

[5] Saurashtrian or Kathiawadi Gujarati refers to the dialects of the Saurashtra region, i.e. the area surrounding Diu; an offshoot spoken in Tamil Nadu, particularly in the region of Madurai, is also identified as Saurashtri or Sourashtri - see Masica (1991:442) and Southworth (2005), but this variety is not relevant for our purposes. South Gujarati is a common designation for the varieties spoken to the east of the Gulf of Cambay. Standard Gujarati, on the other hand, is presently defined as the northern dialect centred around the cities of Vadodara (formerly Baroda) and Amdavad (or Ahmedabad).

agree on the alveolar (which Cardona & Suthar subsume under the general category of dentals) and glottal segments but interpret the third one differently; Mistry calls it a palatal fricative while Cardona & Suthar classify its articulation as postalveolar, which for them corresponds to a retroflex articulation (the notation observed here). It should be noted that /z/ occurs only in words of Arabic, Urdu, Persian or English origin and it is not realised as a voiced segment by all speakers. The same holds true for /ʂ/, which contrasts with /s/ in very few cases and occurs prominently in loanwords; once again, the contrasts only obtain in certain dialects.

Mistry (1997:659) claims that /ʋ/ has two allophonic realisations, [ʋ] proper word-initially and [w] in every other position, whereas both are allowed in between vowels.

9.1.3   The alignment of DIP phonology and phonetics

This section deals with two related questions, labelled as follows:

A - What have Portuguese and Gujarati contributed to the phonological make-up of modern DIP and how?

B - In cases where Modern Portuguese and Modern Gujarati differ from their 16th-century counterparts, has DIP accompanied change in (any of) these languages?

Answering A implies probing the evidence available for B, i.e. in assessing the contributions of Gujarati and Portuguese one must compare DIP with both the present and past stages of these languages; therefore, this comparative study will debate both simultaneously. The phonological system, phonetic principles and phonotactics are all within the scope of this section, and the discussion is segmented in the interest of clarity.

9.1.3.1   The phonological system: vowels

Let us first raise a few points concerning the parentage and innovation of DIP's vocalic system. It is not surprising that DIP maintains a 4-height system for its oral vowels, as this mirrors both Gujarati and Portuguese. The comparison is more complex than this, however, given that on a phonological level DIP mid-high and mid-low vowels do not contrast in unstressed position. In Gujarati, as clarified above, the very opposition is overall rather marginal and confined to certain dialects. In a different way, and in contrast with Classical Portuguese, the opposition is also cancelled in Modern Portuguese in unstressed syllables. DIP did not entirely follow the additional diachronic changes affecting Portuguese unstressed mid vowels (raising and centralisation), and instead optionally retains the differentiated realisations it received from Classical Portuguese input. The partial neutralisation of mid-high/mid-low opposition in DIP may therefore be understood as the result of a combination of factors from both parent languages: Gujarati's weak opposition and the annulment of the contrast in Modern Portuguese in unstressed position. The maintenance of the opposition in stressed position, on the other hand, sides with both

Classical and Modern Portuguese and does not contradict Gujarati precisely because the contrast does exist in this language even if marginally.

As seen in section 5.1.3, DIP does retain Classical Portuguese realisations of mid-high and mid-low segments in unstressed position, even though an optional heightening pull may optionally turn [ɛ]/[e] into [i] and make [ɔ]/[o] surface as [u]. This fact is quite significant for the debate concerning the diachronic development of European Portuguese pre-stress vowels. Nowadays, except for a few exceptional cases, no mid-high and mid-low vowels occur in non-initial pre-stress position in European Portuguese (see Mateus and d'Andrade 2000); Figure 9.1, though, makes it clear that Classical Portuguese maintained a distinctive four-height system in unstressed position (cf. Castro 1991:252), before a widespread wave of unstressed vowel heightening. The exact chronology of this heightening, and in particular of the precise moment in which it affected the entire territory of Portugal, has been a matter of some debate, but the fact that the distinction is relevant for DIP (as well as for Brazilian Portuguese and some of the Gulf of Guinea Creoles) suggests that its *terminus ante quem* cannot be set before 1535 (see section 3.1.2), at least. It is unlikely that the annulment of the unstressed mid-high/mid-low opposition took place in the initial stages of DIP formation *contra* the etymological material, and as such we can postulate that the blurring of boundaries was a subsequent development, arguably spurred by (but not totally tied to) the diachronic change affecting standard Portuguese.

The absence of /ɐ/ from DIP cannot be immediately explained with reference to one or another stage of Portuguese, as this segment occurs in both. The absence of this segment from Gujarati may therefore have been decisive for the fact that it was unretained (or lost) in DIP.

Some of the diphthongs attested in DIP, namely [ɛj] and [ej] are no longer licensed in Modern Portuguese, but according to Clements (1996:63) both were present in Classical Portuguese. On the other hand, [əj] has never featured in SP, whereas it occurs in both DIP (in the word /dəpəj/ 'afterwards') and Gujarati (e.g. /səməj/ 'time'; /pəjsa/ 'money'). Despite this, it is no exaggeration to say that the DIP diphthong system is closer to Portuguese than to Gujarati. DIP significantly contrasts with Gujarati in that it disallows rising diphthongs, with the exception of those following velar plosives; crucially, this exception also applies to Portuguese. The only rising diphthong of DIP that is not shared with Portuguese is instantiated in the 1st person singular pronoun *yo*, but see chapter 5 fn. 9 for a possible diachronic explanation of this exceptional form. Conversely, it has been shown above that DIP does not form the diphthong *[ow]. It is unclear whether this diphthong was ever present or whether monophthongation generalised at some point. However, it is interesting to notice that Portuguese went through a process of monophthongation of [ow] into [o] during its Classical period; the diphthong is still preserved in the northern varieties of European Portuguese, but it could have been the case that the varieties most relevant as sources of DIP lexemes had already completed the process by the time of the initial contact.

There is a fundamental difference between DIP vowels and Gujarati vowels, particularly the absence of breathy voiced variants in the former. The only conclusion we can draw from this fact is that contact with Gujarati was not able to add a trait absent from the Portuguese input. This is not to say, as demonstrated above, that Gujarati had no influence on DIP phonology; we have already seen that, as far as the vocalic system and

sub-systems of DIP are concerned, the contribution of Gujarati operates mostly in the form of 'ganging-up effects' (see e.g. Aboh and Ansaldo 2007) (i.e., the reinforcement of tendencies common to the two ancestor languages). This becomes even more evident as we move on to the consonantal system.

9.1.3.2   The phonological system: consonants

The first observation to be made is that, unlike Gujarati, DIP does not contain retroflex phonemes. This results in significant differences between the two phonological systems affecting plosives, nasals as well as fricatives and laterals.

DIP plosives closely mirror the points of articulation of Portuguese plosives, and they share with both Portuguese and Gujarati the pattern of voiced/voiceless pairing. DIP differs from Gujarati in that it does not contain aspirated plosives. Consider that, according to certain interpretations (see previous section), Gujarati affricates are best interpreted as palatal plosives. I treat them as affricates, which in combination with the fact that affricates also occur in Classical Portuguese (and modern dialectal Portuguese) begins to explain the retention of such segments in modern DIP. This issue is taken up once again below, when we discuss the DIP fricatives.

Both Portuguese and Gujarati have bilabial and alveolar nasal segments, which is mirrored in DIP. The velar nasal /ŋ/ of DIP, however, is an interesting development, as it consistently comes to replace the etymological palatal nasal /ɲ/ (e.g. Ptg. /tiɲɐ/ 'had' for DIP /tiŋ/ 'had'). This change cannot be attributed to Gujarati because, even though the velar segment does occur, it is restricted to a very particular context, viz. between a nasal vowel and a velar plosive. On the other hand, there is evidence from the diachrony of Portuguese to suggest that in the 16th-century input the occurrence of the nasal palatal must have been less solid than in Modern Portuguese. The segment existed in Old Portuguese (most notably as the development of the Latin sequence -gn-), but Teyssier (2001:41) makes it clear that it was just prior to the transition from Old Portuguese to Classical Portuguese (around the 14th century) that most of the instances of this consonant came into being, when it was epenthetically inserted to resolve any hiatus consisting of a nasal vowel (with the exception of /ũ/) plus an oral vowel. This trend led to transformations of the type [ra'ĩa] > [ra'iɲa] 'queen' or [di'̃ejro] > [di'ɲejro] 'money'; an important point is that this development denasalises the first segment of the original hiatus. Castro (1991:216) admits ɲ-epenthesis may have begun as early as the 13th century and claims that the process was complete sometime during the 15th century, but it is unclear whether the written sources available are representative of the entirety of the Portuguese dialects. It is admissible, though by no means provable beyond doubt, that the input available at the initial stages of DIP formation had not consistently finalised this change and that DIP was therefore freer than assumed to deviate from the solution that was to become standard in Portuguese. The sequence present to this day in DIP, viz. a velar approximant with maintenance of nasality on a previous high-front vowel or glide, is distinct from standard Modern Portuguese not only in the mode and point of articulation of the consonant but also in that it retains the original nasality of the vowel. According to this scenario, then, the present correspondence between Ptg. /ɲ/ and DIP

/ŋ/ (phonetically [ɯ]) is not the result of the alteration of an etymological palatal nasal into a velar consonant, but rather the result of two diverging paths of change.

The opposition in Portuguese (both Classical and Modern) between a trill and a flap has no correspondence in DIP, although both sounds do occur in allophonic distribution. In this respect, then, DIP sides more closely with Gujarati. Apart from the phonological flap, in Gujarati an intervocalic nasal can also be flapped, which does not happen in DIP.

The case of the fricatives is somewhat complex. The Gujarati glottal fricative is absent from both DIP and Portuguese. As far as the labiodental point of articulation is concerned, DIP mirrors the phonological segments of Portuguese, but it must be noted that the realisation of DIP /v/ as [ʋ] is the result of Gujarati influence. Both Gujarati and Portuguese have voiced and voiceless alveolar fricatives /s/ and /z/, and that is also the case in DIP. It has already been shown that two particular phonemes (here represented as /tʃ/ and /dʒ/) may optionally trigger an affricate as well as a fricative realisation. It is clear that, in these cases, the affricate was the original realisation of the phoneme in the earlier stages of the language, as this was still a characteristic of European Portuguese in the early 16th-century and, furthermore, one that was preserved in a number of Portuguese-lexified Creoles as well as Portuguese loans in several South Asian languages.[6]

In Old Portuguese, up to the 16th century, the affricate [tʃ] [graphed <ch> in old documents] is said to have contrasted with [ʃ] (represented by <x> in the same sources), and the two only fused into /ʃ/ from the 17th century onwards.[7] With respect to the voiceless fricative/affricate divide, then, DIP closely mirrors the system that was prevalent in early Classical Portuguese. But the parallel is all the more striking when we realise that, just like DIP, the contrast between the postalveolar fricative and the affricate was limited in Classical Portuguese to the voiceless segments, i.e. no phonemic opposition ever obtained between [dʒ] and [ʒ] in Portuguese.[8]

Some time during the Classical Portuguese period, syllable-final /s/ came to be pronounced [ʃ] or [ʒ] (the voice feature dependent on the following segment). This change was not fully accompanied by DIP, in which [s] still occurs syllable-finally but is sometimes replaced by [ʃ], giving rise to variation patterns of the type [kas'tɛl] vs. [kaʃ'tɛl] 'fort', [is'kɔl] vs. [iʃ'kɔl] 'school' (cf. section 5.2.4). Once again, DIP mirrors a development in standard Portuguese without compromising the earliest DIP strategy.

---

[6] Saramaccan, for instance, retains the affricates in items such as *bítju* 'worm' (from Ptg. *bicho* 'worm,animal') and *djái* 'garden' (from Ptg. *jardim* 'garden'); see for example Cardoso (2003). Cape Verdean Creole retains the affricate segments in words such as *sibitxi* 'rosary' [from Ptg. *azeviche* 'jet'], as does Guinea-Bissau Creole in words such as *batxarel* 'loudmouth' [from Ptg. *bacharel* 'bachelor'] (Rougé 2004). Other varieties of IP also retain the affricate series, such as that of Cochin (e.g. *djenti* 'people', own recording), and the same holds for other Portuguese-lexified Asian creoles, as exemplified by Papia Kristang (e.g. *ntʃídu* 'full', v. Hancock 1975:224). Among the common loans in South Asian languages to preserve the affricate sounds are the cognates of Ptg. *chave* 'key' (*cāvi̇̄* in Marathi, *cābhi̇̄* in Hindi) and *janela* 'window' (*janēlaya* in Sinhala, *jānālā* in Bengali, *janglā* in Hindi) - Masica (see 1991:73-74).

[7] The process is incomplete in the Portuguese territory, as certain dialects have not yet fused the two.

[8] Teyssier (2001:27) argues that this phoneme originally had an affricate pronunciation and eventually changed into a postalveolar fricative, and claims that it is unclear whether this change occurred during the Old Portuguese period (i.e., up to the 16th century) or later. The evidence now unearthed from DIP, as well as that provided by Portuguese loans in a number of South Asian languages and other creoles (see this chapter fn. 6), clarifies that the earlier affricate realisation must have been dominant in Portuguese at least until the 1530s.

If we consider Gujarati /ʋ/ to correspond to Portuguese /w/, we will notice that the two languages are equal with respect to the number of phonological glides and very similar as to their points of articulation. DIP accordingly has two phonological approximants /w/ and /j/, although on the phonetic level an extra approximant /ʋ/ occurs. According to my interpretation of the relationship between DIP [ʋ] and [w] in 5.2.4, contact with Gujarati was able to affect the realisation of DIP /v/ but did not proceed to undo the opposition between /v/ and /w/. However, as I also admit there, this interpretation is open to debate, in which case one would have to admit a more radical influence of Gujarati onto the subsystems of DIP fricatives and approximants.

DIP retains only one of the lateral segments provided in the combined input of Gujarati and Portuguese, namely /l/. Crucially, this is the only segment shared by both ancestor languages (although in DIP it tends towards a dental instead of alveolar realisation), which suggests the operation of 'ganging-up effects' by which traits shared by all contributors to the original feature pool are particularly likely to be selected into the budding contact variety.

9.1.3.3 Phonetic processes

Whereas the influence of Gujarati on DIP phonology has been somewhat modest, the language takes up a prominent role when we turn to analysing phonetic parameters. Most of the conspicuous phonetic processes of DIP described in chapter 5 have clear counterparts in Gujarati. These include:

a) unreleased final plosives: in Gujarati, as in DIP but not in Portuguese, plosives can occur in word-final position; in these positions they can be either released or unreleased: e.g. [kəd] vs. [kəd̚] 'size' (Mistry 1997:660), cf. DIP [pɛd̚] 'stone'. When a plosive occurs in syllable-final position and the subsequent segment is consonantal, it is always unreleased (e.g. [ək̚ʂər] 'syllable', adapted from Mistry 1997:660);

b) schwa-deletion: the process of schwa-deletion in DIP was characterised as affecting middle syllables of trisyllabic words; see sections 5.1.3 and 5.4. In Gujarati, schwas are often deleted in -CəCV# contexts: e.g. /pələʈo/ 'change [IMP]' is realised [pəlʈo] (adapted from Mistry 1997:661), although the process may be suspended if the word ending corresponds to certain grammatical markers;

c) non-simplification of geminates: like DIP, Gujarati admits geminates. One important difference is that while in DIP geminates only arise in discourse and never occur within a word, in Gujarati consonant geminates contrast with single consonants: e.g. /muko/ 'put [IMP]' vs. /mukko/ 'fist' (adapted from Mistry 1997:670); a consonant can also be duplicated for intensification: e.g. /nəʋũ/ 'new' vs. /nəʋʋũ/ 'quite new' (*idem*);

d) incomplete constriction of labiodental segment: as described above, Gujarati does not contain the segment [v], but it does contain an approximant [ʋ]. This is a trait shared with several other Indo-Aryan languages and must be seen as the driving force behind the

realisation of DIP /v/ as [ʋ].

9.1.3.4 Phonotactics

In order to explore the phonotactic links between DIP and both of its ancestor languages, I will contrast these language's consonant clusters, concentrating on complex syllable onsets alone because information is more readily available for Gujarati complex onsets than for syllable-final clusters.

The structure of complex syllable onsets in DIP, as described in 5.3.1, is such that the second segment in consonant clusters must be either /r/ or /l/, and the first must be a plosive or a fricative. In this respect DIP mirrors Portuguese much more closely than Gujarati. Portuguese onset clusters overwhelming admit only /ɾ/ and /l/ as their second element after either a plosive or a fricative (e.g. *primo* 'cousin [m.]', *claro* 'clear', *fresco* 'fresh, fresco'). Other clusters such as *pn* (e.g. *pneu* 'tyre') or *gn* (e.g. *gnomo* 'gnome') only occur in learned loans, and because of their violation of the Sonority Principle it has been proposed e.g. Mateus and d'Andrade 2000:40*ff* that they are in fact broken up at the phonemic level by an empty nucleus that does not surface in European Portuguese (e.g. *pneu* is pronounced [pnew]) but does in Brazilian Portuguese (in which *pneu* is pronounced [pi'new]). Gujarati, on the other hand, contains a number of consonant sequences unrepresented in DIP. Complex onsets in Gujarati are usually occupied by:[9] a) a plosive followed by /r/, /l/, /j/ or /ʋ/ (e.g. *krupa* 'kindness', *kvath* ''decoction'); b) /m/, /n/, /l/ or /ʋ/ followed by /j/ (e.g. *nyay* 'justice', *lyanət* 'disgrace'); c) /m/ or /ʋ/ followed by /r/ (e.g. *mrutyu* 'death', *vrət*) 'oath'; d) /m/ followed by /l/ (e.g. *mlan* 'withered'); e) /s/ followed by /j/, /ʋ/, /r/, /k/, /t/, /ʈ/, /n/ or /m/ (e.g. *svikar* 'accepting', *smarək* 'memorial'); and f) /ʃ/ followed by /j/, /ʋ/, /r/, /l/ and /m/ (e.g. *švas* 'breath', *šməšan* 'cremation ground'). The language furthermore contains some triconsonantal clusters in initial position, in Sanskrit, Arabic, Persian and English loans (see Cardona 1965).

Finally, a comparison of the three languages' stress assignment strategies reveals DIP departs in very fundamental ways from both Gujarati and Portuguese. DIP stress falls obligatorily on the final syllable, and certain function words appear not to be assigned stress at all (see section 5.4). Stress on content words is therefore phonologically (rather than lexically) assigned. This establishes an important parallel with Gujarati in that stress is not contrastive, although Gujarati follows quite different stress assignment rules. Despite theoretical disagreement among linguists, a powerful generalisation predicts that, in Gujarati: a) disyllabic words have word initial stress unless the first syllable contains a vowel other than /a/ and the final syllable contains /a/; b) the first syllable of trisyllabic words is stressed unless the first syllable contains /a/, in which case speakers vary as to whether they place stress on the first or second syllable (see Cardona and Suthar 2003:666). The source of disagreement in this matter probably stems from the fact that, as noticed by Turner (1921, 1973), stress in Gujarati is not very pronounced. DIP somewhat mirrors this reality in the sense that, whereas stress-assignment regularities are recognisable, sandhi and discourse features may overrule them so that non-final syllables can in fact supersede the final one in pitch and intensity - cf. section 5.4 as well as chapter 5 fn. 12.

---

[9]All examples take from Cardona and Suthar (2003), and graphed accordingly.

In Portuguese, by contrast, stress is distinctive, as shown by the pair *júbilo* (pronounced [ˈʒuβilu]) 'rejoicing (n.)' vs. *jubilo* (pronounced [ʒuˈβilu]) '(I) rejoice'. Any generalisations fail to predict the assignment of stress, and there are different general rules for nominal and verbal roots (see Mateus and d'Andrade 2000:109*ff*). One general constraint does apply to the effect that stress must be placed on one of the three last syllables.

9.1.3.5   Summary

This comparative survey reveals a number of important facts concerning the relative influence of both Gujarati and Portuguese on DIP phonology and phonetics, which are summarised below:

- On the phonological level, DIP tends to follow the phonemic distinctions inherited from its main lexifier, Portuguese, with some alterations (e.g. DIP /ŋ for Portuguese /ɲ/) and rare suppression (e.g. the suppression of Ptg. /ʎ/). Gujarati has not been able to add any features to DIP that are absent from Portuguese (e.g. breathy voice, aspiration, retroflexion) but it has contributed to subtract from the Portuguese input (e.g. the suppression of Portuguese trills /r/ and /ʀ/);

- Wherever phonological change affected Portuguese from the 16th century onwards, DIP often retains the features of Classical Portuguese (e.g. the maintenance of affricates), but it may also adopt a Portuguese innovation without abandoning its original feature, therefore resulting in variation (e.g. the alveolar vs. palatal realisation of syllable-final fricatives);

- On the level of phonetic realisations, in contrast with what happens on a phonological level, the influence of Gujarati is very notorious (e.g. schwa-deletion, the unreleased nature of syllable-final plosives).

9.2   Morphology

DIP establishes considerably less morphological oppositions than either Gujarati or Portuguese. To illustrate the extent to which DIP is at odds with the morphological requirements of both these languages, I will contrast both nominal (section 9.2.1) and verbal morphology (section 9.2.2) for the three languages.

9.2.1   Nominal morphology

Gujarati features variable and invariable nominal stems. For variable stems, the language makes use of nominal inflection to code gender (masculine, feminine and neuter) and number (singular and plural). Masculine, feminine and neuter gender are marked by the suffixes *-o*, *-i* and *-ũ* respectively. The suffix *-o* indicates plural reference, although it is

optional in case any other element of the utterance clarifies the plural reference. Consider the following examples, from Cardona and Suthar (2003:670):

(336)   a.   *rameš-e cɔpḍ-i-o khǝrid-i.*
R.-AG   book-f-p buy-PFV.f
'Ramesh bought some books.'

   b.   *rameš-e trǝṇ   cɔpḍ-i khǝrid-i.*
R.-AG   three book-f buy-PFV.f
'Ramesh bought three books.'

The presence of a numeral in (336b) cancels the need for overt plural marking on *cɔpḍi*. But these examples reveal yet another domain of nominal modification in Gujarati, namely the postposed clitic elements indicating the syntactic/semantic role of a given NP; notice the occurrence of Agentive *-e* in (336). Whenever followed by one of these postposed elements, the nominal form may occur either in its base form (preceding the agentive clitic *-e* or the locative clitic *-e*) or an oblique stem with a different set of gender/number suffixes (feminine nouns have less morphological variables than masculine or neuter nouns). In the case of the neuter noun *ǝṭhvaḍiũ* 'week', the base form selected is *ǝṭhvaḍi-* and the oblique form takes the suffix *-a*, resulting in *ǝṭhvaḍia*; their occurrence is exemplified in (337), adapted from Cardona (1965):

(337)   a.   *awte ǝṭhvaḍi-e.*
other week-LOC
'next week.'

   b.   *hũ vǝḍodra-mã   ek   ǝṭhvaḍia-thi chũ.*
1s Baroda-LOC one week-ABL   AUX.1s
'I have been in Baroda for one week.'

Invariable nouns, which do not admit the full array of suffixes described above, are nonetheless assigned a gender value, grammatical or natural, which is revealed in the agreement patterns of other sentential constituents; for instance, *hoṭh* 'lip' is masculine, *jibh* 'tongue' is feminine, and *mŭch* 'moustache' is neuter.

As far as Portuguese is concerned, a great deal of grammatical information is also overtly coded through nominal morphology. The form of nouns provides clear indication of gender (feminine or masculine) and number (singular or plural), with the exception of a few invariable forms. All nouns have intrinsic gender, be it natural (e.g. *gat-o* 'cat [m.]', *gat-a* 'cat [f.]') or purely grammatical (e.g. *oboé* 'oboe [m.]', *flauta* 'flute [f.]'). Variable nouns accept both a feminine and masculine suffix to indicate natural gender. Typically, *-a* functions as the feminine suffix and *-o* marks the masculine, but other markers also occur (e.g. *impera-dor* 'emperor [f.]', *impera-triz* 'empress [f.]'). Invariable nouns do not admit this type of alternation, although they are still assigned gender; these may be masculine,

feminine (e.g. *homem* 'man [m.]', *mulher* 'woman [f.]') or both (e.g. *o mártir* 'the[m.] martyr', *a mártir* 'the[f.] martyr', in which the gender variable is revealed in the form of the article).

Plural reference is obligatorily marked on the relevant noun (and other clausal constituents, through agreement). Notice example (338), in which the selection of a numeral (*três* 'three') does not impede number marking on the head of the NP and the definite article. Notice also how the form of both the article and demonstrative reflect the gender (feminine, in this case) associated with the nouns they modify.

(338) *aquel-a raposa come-u   a-s   três   galinha-s.*
       DEM-f fox.f   eat-PST.3s ART.f-p three chicken.f-p
       'That fox ate the three chickens.'

In general terms, the absence of a number suffix indicates singular while the suffix -*s* indicates plural; the realisation of the plural suffix may at times entail formal changes to the root (e.g. *pombal* 'dovecote', *pombai-s* 'dovecotes'). As expected, mass nouns such as *leite* 'milk [m.]' do not normally admit plural marking. The exceptions to the general rules can usually be explained diachronically.

In stark contrast with the nominal inflection strategies of Gujarati and Portuguese, DIP does not mark number, gender or case by morphological means (see section 6.4). The language does make use of an optional marker of plurality (interpreted as a collectiviser in 7.2.7) *tud*, equivalent to the universal quantifier meaning 'all'. This element is preposed to the noun to indicate additive plurality and postposed to construct similative plurality - in which the members of the group are identified as sharing some trait with an expressed referent (see 7.2.7.2).

Various Indian languages obtain similative plural reference through a morphological phenomenon known as *echo word formation* (see e.g. Emeneau 1956, Masica 1991). This consists of the reduplication of a noun in which the initial syllable or consonant of the second occurrence of the word is replaced with a predefined segment. In Gujarati, the initial consonant of the repeated word is replaced with *b* or *ph*, and these segments are inserted in the case of a vowel-initial word, for instance *chətri-bətri* 'umbrellas and such', *orḍo-borḍo* 'rooms and such' (Cardona and Suthar 2003:679).[10]

The form of the similative plural in DIP is rather different from an echo-word construction, and as such one cannot recognise the direct formal influence of Gujarati in this matter. However, the prevalence of similative plural reference in Gujarati (and other South Asian languages) is likely to have provided the pragmatic motivation for DIP to grammaticalise a semantically equivalent construction.

---

[10] This structure is indicated for Dravidian languages by Masica (reported in Daniel and Moravcsik 2005), who provides the following Telugu example: *puli gili* 'tigers and such' (from *puli* 'tiger') and Subbarao (2008), who adds the expression *ceruku giruku* 'sugarcane and the like' (from *ceruku* 'sugarcane'). Subbarao (2008) also reports on Hindi/Urdu echo words such as *caay waay* 'tea and the like' (from *caay* 'tea') and *kʰaanaa waanaa* 'food and the like' (from *kʰaanaa* 'food').

9.2.2   Verbal morphology

Most Portuguese verbs participate of one of three regular conjugational classes which to a large extent determine the form of the adequate suffixes. Other verbs fall outside these conjugational classes, constituting irregular paradigms. In the extreme, a verbal form will encode TAM categories, person (1st, 2nd, 3rd) and number (singular, plural) by morphological means. This subsystem is highly fusional in that verbal suffixes often collapse temporal, modal and/or aspectual categories, on the one hand, and both number and person, on the other. Consider the following example, the segmented analysis of the subjunctive verb form *abandonássemos* in Portuguese:

(339)   *abandon-á-sse-mos*
        abandon-CLS$_1$-SBJV.PST-1p
        '[that we] abandoned'

Suppletion is also common in Portuguese; consider for instance the following examples, which differ only with respect to TAM categories:

(340)   a.   *vás*
             go.SBJV.PRS.2s
             '[that you] go'

        b.   *fo-ste*
             go.IND.PST-2s
             '[you] went'

        c.   *ir-ás*
             go.FUT-IND.2s
             '[you] will go'

Portuguese imperative verb forms are formally unique in that they follow specific inflectional patterns which distinguish them from other verb forms.

In Gujarati, the verbal form also codes TAM categories, person (1st, 2nd, 3rd) and number (singular, plural). Lexical verbs code person and number categories by affixation, (although certain cases of suppletive marking do occur, cf. the base form of the auxiliary verb: *ch-* in the present, *hɔ-* elsewhere) and, in the absence of an argument with which to agree, select the neuter suffix -*ŭ*; cf. example (341c). Some examples follow:

(341)   a.   *av-ŭ̃*
             come-1sg
             '[I] come'

b. *av-š-e*
    come-FUT-3
    '[he/she/it] will come'

c. *av-y-ũ*
    come-PFV-n
    '[you] went'

As shown in (341b,c), future tense is indicated on the verb by -(*i*)*ʃ*-, while the affix -*y*-
expresses perfective aspect. The imperfective affix is -*t*-. Consider the examples in (342):[11]

(342) a. *kišor dhime bol-y-o.*
           K.   low   speak-PFV-m.s
           'Kishor spoke softly.'

       b. *te   jəman-e hũ dər əṭhwadi-e vədodra jə-t-o.*
           DEM time-LOC 1s every week-LOC Vadodara go-IPFV-m.s
           'At that time, I would go to Vadodara (Baroda) every week.'

       b. *huN kale   təmahre tyaN aw-iš.*
           1s   tomorrow POSS.2 home come-FUT
           'Tomorrow I will come to your home.'

The base form of the verb is used as an imperative (e.g. *av* 'come!'), to which a polite
plural suffix -*o* can be appended (e.g. *av-o* 'please come!'). The future imperative is
characterised by the infix -*j*-, followed by the ending -*e* (e.g. *av-j-e* 'you will come!') or
polite plural -*o* (e.g. *av-j-o* 'you will please come!').[12]
   A special affix codes conditional mood (343a). In addition, Gujarati also features a
morphological causative (343b) and passive (343c), constructed with resort to affixes[13]
(from Cardona and Suthar 2003):

(343) a. *av-ət*
           come-COND
           'if [X] come(s)'/'[X] would come'

---

[11] Examples (342a,b) adapted from Cardona and Suthar (2003:682*ff*). The Future tense affix is graphed
-*iš* in (342c), in accordance with the original, in Cardona (1965:104).

[12] In addition, the infinitive form of the verb may also constitute an imperative; in such cases, the absence
of honorific 2nd person inflection reveals that it is less respectful and more insistent. As described in section
7.6.3.1, this is also true of the use of infinitive as an imperative in DIP. This possibility in DIP is clearly
modelled on Gujarati, given that Portuguese does not allow an infinitive form in a requestative speech act.

[13] See 9.6.4 for a description of Gujarati passive constructions.

b. *ɛ-ne   bɛs-aḍ-o.*
3s-OBJ sit-CAUS-2p
'Please seat him.'

c. *əNgreji-mã   manṇəs-ne šũ   kəhv-a-y?*
English-LOC man-OBJ   what say-PASS-3p
'How do you say "man" in English?' [Lit. 'What is "man" said in English?']

When compared to both Gujarati and Portuguese, then, DIP verbal morphology appears rather reduced. In this language, no person, number and modal categories are expressed morphologically. Only perfective forms are regularly inflected, coding Past and non-Past tense. Furthermore, there is an infinitive suffix and, in the speech of some speakers, a participle and gerund forms - see 6.2 for further information.

9.3   Paradigms

The paradigms of personal pronouns, possessives and deictic elements in Gujarati and Portuguese have important points of contact but also important differences. I will therefore carry out a comparative study in order to ascertain to what extent DIP can be said to align with one or the other of its ancestors, and to what extent (and by what means) it departs from both. Given that both personal pronoun and possessive paradigms ode person distinctions in all three languages, they will be discussed simultaneously in section 9.3.1; section 9.3.2 is concerned with a specific characteristic of deictic systems, viz. the conceptualisation (and linguistic manifestation) of the deictic space.

9.3.1   Personal pronouns and possessives

The form of Portuguese personal pronouns and possessives varies to express person (1st, 2nd, 3rd) and number (singular, plural). There are some distinctions between the Classical and Modern Portuguese paradigms, which are captured in Table 9.4.[14]

Personal pron.		*Person*	*Possessives*	
SG	*PL*		*SG*	*PL*
eu, me[P,U,R], mim[OBL]	nós, nos[P,U,R]	*1*	meu, minha	nosso(s), nossa(s)
tu, te[P,U,R], ti/si[OBL]	vós, vocês	*2*	teu(s), tua(s)	vosso(s), vossa(s)
você, lhe[R] [f. late 16C]	vos[P,U,R], lhes[R]		seu(s), sua(s) [f. late 16C]	
ele[m.], ela[f.]	eles[m.], elas[f.]	*3*	seu(s), sua(s)	seu(s), sua(s)
o/a[P,U], lhe[R]	os/as [P,U], lhes[R]		(dele[m.], dela[f.])	(deles[m.], delas[m.])

Table 9.4: Portuguese personal pronouns and possessives

---

[14]See section 7.5 for an account of the abbreviations P, U and R, which stand for different argument roles in various types of clause.

Only the 1s and 2s personal pronouns have specific oblique forms (*mim* and *ti/si*, the latter being the polite form), but for all pronouns there are 'object' forms (often cliticised), a split between Recipient forms and Patient/Undergoer forms applying only to 3rd person. It is important to notice that the use of *lhe* and *lhes* as 2nd person (clitic) pronouns is an extension from the 3rd person forms accompanying the tendency to use a 3rd person form of address to engage a high-ranking interlocutor; an old such form of address was *vossa mercê* 'your mercy', the erosion of which into *você* was only stabilised in the late 16th century (see Teyssier 2001). Gender distinctions (masculine, feminine) are only coded in 3rd person personal pronouns.

Possessives are NP operators and, as such, agree in number and gender with the nominal head. It must be noted that, at present, the 3rd person forms *seu(s)/sua(s)* compete with different compound forms involving the relator preposition *de* 'of' and the personal pronouns, which however behave rather differently from the original possessives: *dele/(s)*, *dela/(s)* typically occur postnominally (unlike prototypical noun modifiers) and their gender/number characteristics depend on their referent instead of the noun they modify.

Gujarati personal pronouns and possessives are given in Table 9.5:

Personal pron.		*Person*		*Possessives*	
SG	*PL*			*SG*	*PL*
hũ	apṇ-e[INCL], əm-e[EXCL]	*1SBJ*			
mə-ne	apṇ-ne[INCL], əm-ne[EXCL]	*1OBJ*	*1*	mar-	amar-
mə̄	apṇ-e[INCL], am-e[EXCL]	*1AG*			
tū	təm-e, ap[formal]	*2SBJ*			
tə-ne	təm-ne, ap-ne[formal]	*2OBJ*	*2*	tar-	təmar-
tə̄	təm-e, ap-e[formal]	*2AG*			
e, te	teo, e	*3*		ten-	teon-, temn-
	[Equivalent to deictic pronouns]				[Deictic pron. + relator -*n*-]

Table 9.5: Gujarati personal pronouns and possessives

When approaching the personal pronoun paradigm of Gujarati, it is convenient to separate the forms according to their various syntactic and semantic functions because, even though the bulk of the distinction is encoded through the suffixes, these involve some root alternations as well. Gujarati codes an inclusive/exclusive distinction in 1st person pronouns. As far as 2nd person pronouns are concerned, it must be noted that often the singular form is used with familiar interlocutors while *təm-* and *ap(-)* in formal and very formal situations respectively, irrespective of the actual singularity/plurality of the interlocutor(s) (see Tisdall 1892, Cardona and Suthar 2003). Gujarati has no dedicated 3rd person personal pronoun forms, instead making use of deictic (demonstrative) elements - see 9.3.2.

In a sense, one could say that Gujarati possessives are simply genitive-marked forms of the personal pronouns (which is particularly evident in the case of the 3rd person forms), but in 1st and 2nd person the base form is suppletive. Possessives admit gender and number suffixes similar to those marking nouns (and other noun modifiers). Possessives precede the noun they modify and agree with it in number and gender.

The DIP personal pronoun and possessive paradigms are repeated in Table 9.6 for

easier comparison (see also sections 6.5.4 and 6.6.1).

Personal pron.		Person	Possessives	
SG	PL		SG	PL
yo, mi[OBL]	nɔs	1	mi	nɔs
use	usez	2	duse	dusez
el[m.], ɛl[f.]	e(l)z	3	del[m.], dɛl[f.]	de(l)z

Table 9.6: Diu Indo-Portuguese personal pronouns and possessives

The forms of the DIP personal pronouns are clearly derived from Portuguese, but with some important differences: a) only the 1st person pronoun retains an oblique form (used, for instance, in combination with a preposition); b) no 'object' forms exist;[15] and c) the gender distinction in 2rd person plural has been suppressed. 2nd and 3rd person pronominal forms retain remnants of the Portuguese plural suffix *-s* which, however, is not productive in DIP. It is also curious to notice that the 2nd person forms derive from the Portuguese *você*, i.e., the form which was still not entirely widespread in Portuguese by the time of the formation of DIP, instead of the earlier (and still active) forms *tu* and *vós*.[16] In contrast with Gujarati, the 1st person personal pronouns in DIP do not code an inclusive/exclusive opposition.

The possessive forms in DIP reveal a greater amount of innovation in relation to the Portuguese paradigm. The *de* + Personal Pronoun structure responsible for Portuguese *dele/(s)*, *dela/(s)* is mirrored in the DIP 3rd person possessives forms *del*, *dɛl* and *de(l)z* but extended to the 2nd person forms *duse* and *dusez*. As explained in 7.2.3, there are records of the pattern being further extended to include 1st person possessives, although this is not the norm. Finally, the form of the 1st person singular possessive *mĩ* is equivalent to the oblique form of the 1s personal pronoun.[17]

9.3.2 Deictics

In truth, personal pronouns and possessives (discussed in 9.3.1) are deictic elements in that their correct reference can only be apprehended with reference to (discourse) context. In this section, I will simply contrast DIP, Gujarati and Portuguese with respect to the organisation of deictic space manifested in demonstratives and certain spatial adverbs.

---

[15] In the case of the 1st person singular, the oblique form takes up this function, given that all pronominal arguments in P, U or R position are obligatorily dative-marked; the issue is described in full in section 7.5.

[16] This fact is all the more striking if we consider that the corresponding form in Daman IP, a close variety of IP with a similar formative history, is ɔs, which appears to derive from Portuguese *vós*. DIP *use* was, however, not the original form in DIP. In the 19th-century DIP corpus published in Schuchardt (1883), the only 2p personal pronoun form recorded is *ós*; we are dealing with an instance in which DIP was influenced in recent times by input from standard Portuguese but Daman IP was not.

[17] This may be coincidental given that the Portuguese paradigm contains the feminine form *minha*. According to the general adaptation rules observed elsewhere (viz. the omission of etymological posttonic vocalic material, the treatment of etymological /ɲ/ as /ŋ/), this form would be selected by DIP as /miŋ/, whose realisation is perceptually very close to /mĩ/.

The deictic space in Portuguese is organised with reference to the speech participants in a tripartite way, which manifests itself in deictic spatial adverbs and demonstratives:[18]

1. the location of the speaker - *aqui* 'here' / *cá* 'here' / *este* 'this' / *isto* 'this'
2. the location of the interlocutor - *aí* 'there' / *esse* 'that' / *isso* 'that'
3. a location alien to both - *ali* 'there' / *lá* 'there' / *além* 'there' / *aquele* 'that' / *aquilo* 'that'

The demonstratives which occur in attributive function include: *este(s)/esta(s)* 'that/those (there)' (agreeing in gender and number with the head of the NP) modify a referent which is perceived to be in the vicinity of the speaker; *esse(s)/essa(s)* 'that/those (there)', placing the referent away from the speaker and in the vicinity of the interlocutor; *aquele(s)/aquela(s)* 'that/those (over there)', which locate the referent away from both the speaker and the interlocutor. The forms *isto*, *isso* and *aquilo* (which are invariable) are exclusively pronominal and reserved for non-humans.

Several noun modifiers in Gujarati are described as deictic elements. These include *a* 'this', *pel-* 'that' (the only one which inflects for gender and number in agreement with the head of the NP), as well as *e* and *te* - these two are used as 3rd person personal pronouns (see Table 9.5) and, as to their deictic value, are variously interpreted as *neuter* (see e.g. Cardona and Suthar 2003) or *distal* (e.g. Tisdall 1892).[19] The spatial adverbs in question are *əhĩ* 'here' and *tyã* 'there'. It is therefore clear that, as far as the conceptual organisation of deictic space is concerned, Gujarati differs substantially from the Portuguese system.

The demonstratives in DIP side with Gujarati in this respect, in that they do not manifest three deictic points but two. Therefore, *es* 'this' is a proximal demonstrative and *ikəl* 'that' is distal (see section 6.5.4). Likewise, *aki* 'here' is proximal and *ali* 'there' is distal.[20] The difference between the DIP and the Portuguese system can indeed reveal Gujarati influence, but it also raises questions regarding the input provided by Portuguese speakers during the formative stages of DIP. According to (Teyssier 2001:67), late 16th-century documents reveal that Portuguese already observed the three degrees of deixis consistently, but prior to that the chronology of the system is unclear.

9.4   Word order

Basic word order is one domain in which there are sharp differences between the input provided by Portuguese and Gujarati. The basic word order of simple declaratives (see

---

[18] The series containing *cá* 'here' and *lá* 'there' is less specific as to exact location and does not reflect this three-tiered concept of deictic space.

[19] Tisdall (1892) mentions some additional inflected 'demonstratives': *oly-* 'that' and *falān-* 'a certain'.

[20] The fact that *ali* covers the space in the vicinity of the interlocutor, in addition to space removed from both interlocutors, is clear from the following command, directed at a child by his mother:

(1)   *say*   *d-ali.*
      leave.NPST   of-there
      'Get out of there.'

9.6.1 for the syntax of content questions) in both Portuguese and DIP is S-V in intransitive clauses, A-V-P in monotransitive clauses and A-V-U-R in ditransitive clauses. Gujarati, on the other hand, is a verb-final language; the basic word order in Gujarati is A-V in intransitive clauses, A-P-V in monotransitive clauses and A-R-U-V in ditransitive clauses.

In terms of basic word order, then, DIP aligns quite clearly with Portuguese, but this does not describe the complete picture. One important aspect concerns the workings of a focus position. In Gujarati, the position immediately preceding the VP is said to be the focus position (where for instance question words occur, see section 9.6.1). This fact has direct correspondence in DIP, as clarified in section 7.8. In Portuguese, on the other hand, focus depends less heavily on the syntactic placement of constituents.

Attributive adjectives precede the noun in both DIP and Gujarati. This is an areal feature common to most languages in the South Asian linguistic are (cf. Subbarao 2008). In Portuguese, on the other hand, they can either be preposed or postposed to the noun.

Adpositional elements are strictly prepositional in both DIP (for some marginal exceptions, see 7.4.8 and 7.4.10) and Portuguese. In Gujarati, the equivalent elements are postpositional (often cliticised). The examples in (344) clarify the different nature of adpositions (enclosed in brackets in the example sentences) in DIP/Portuguese and Gujarati:

(344) DIP [*ku*] *kuyɛr* // [*nə*] *bastãt paiz*
INS spoon   LOC much country
'with (a) spoon' // 'in several countries'

Ptg. [*com*] *colher* // [*em*] *vári-o-s*   *país-es*
INS   spoon   LOC several-m-p country-p
'with (a) spoon' // 'in several countries'

Guj. *camci-*[*thi*] // *juda∼juda*   *des-o-*[*mã*]
spoon-INS   different∼different country-p-LOC
'with (a) spoon' // 'in several countries'

Both Portuguese and Gujarati permit the ellipsis of clausal constituents, but the latter places fewer restrictions on which elements may be omitted. In common with the Indo-Aryan languages in general (see e.g. Masica 1991:396-401), Gujarati allows the null realisation of virtually every constituent: adjuncts, verbs and NPs are all equally eligible for deletion. As described in section 7.8.2, ellipsis in DIP is equally unconstrained, so in this respect the language appears to side with Gujarati.

9.5   Case-marking

Case-marking is one domain in which DIP establishes important parallels with Gujarati, although the former attributes case with resort to prepositions (but see 9.5.3) while the latter does so by suffixation and postpositions. Despite these formal differences - and the

fact that the form of DIP prepositions is derived from Portuguese - there is often a one-to-one distributional correspondence between a DIP case-marker and a Gujarati case-marker. In this section, I will explore only some aspects of case-marking in these languages: dative assignment onto arguments (in section 9.5.1), case-marking in possessive constructions (in section 9.5.2) and, finally, the formal parallels between DIP complex prepositions and related constructs in Gujarati (in section 9.5.3).

9.5.1 Dative assignment

The dative case-marker in Gujarati is the suffix -ne,[21] and its assignment is to a large extent semantically-determined. Cardona (1965) interprets this suffix as an object-marker ('dative-accusative' in other sources) precisely because animate P, U and R arguments all require -ne. This is however not the case for inanimate arguments, which take no case affix at all.

The distribution of DIP dative prepositions a and pə (see section 7.4.1) partly reproduces this basic alignment pattern in Gujarati, with the difference that R is obligatorily dative-marked in DIP whether or not the argument is inanimate (which is also a characteristic of Portuguese, see below); cf. also 7.5 on DIP alignment strategies. Like Gujarati, DIP non-pronominal P and U arguments are only dative-marked if they are animate.

Case-marking on Gujarati S and A arguments reveals a number of intricate semantic distinctions, most of which have a clear parallel in DIP. In Gujarati, if the subject has no control over the action, it is marked with -ne (Cardona 1965, Mistry 2004). 'Dative subjects' occur not only as experiencers of physical sensations but also of mental and psychological states, predicates of necessity, obligation or volitives. Some examples follow:[22]

(345) a. *məh-ne daNt-maN dukhe ch-e.*
1s-DAT tooth-LOC pain EXS-3s
'I have a toothache.'

b. *ma-ne te bābat-mã ras nathī̄.*
1s-DAT DEM matter-LOC interest NEG.EXS
'I have no interest in the matter.'

c. *tam-ne ʃũ joīē ch-e?*
2s-DAT what be.necessary AUX.PRS-3s
'What do you want/need?'

d. *kišor-ne khub kašTa tha-y-ũ.*
K.-DAT much pain happen-PFV-n

---

[21] Not to be mistaken for -n-, a relational (genitive) element which is followed by gender and number markers (see Cardona and Suthar 2003:677). See also section 9.5.3.

[22] The example in (345a) is from Cardona (1965:110), (345b) is taken from Dhruva (1920:143), while (345c) is given in Masica (1991:348) and (345d) in Mistry (2004:7). I retain the authors' original orthography, but add morpheme breaks wherever necessary.

'Kishor was greatly pained.'

Like in Gujarati, S and A arguments in DIP are dative-marked in similar circumstances. Examples (346) repeat some previously discussed sentences for proper comparison:

(346) a. *a mĩ təme sẽt asĩ kom vɛrgoŋ purke nɔs ẽ kaz fal*
DAT 1s.OBL also feel.NPST thus like shame because 1p in house speak
asĩ nə?
thus REQ
'I also feel a bit of shame because at home we speak like this, you see?'

b. *pə ɛl təme aprende-w purtɛgez ku nɔs.*
DAT 3s also learn-PST Portuguese COM 1p
'She also learnt Portuguese from us.'

c. *a el te med.*
DAT 3s have.NPST fear
'He is scared.'

'Dative subjects' are a shared feature of many South Asian languages, commonly given as support for the South Asian linguistic area (see e.g. Masica 1976, Subbarao 2008). In Portuguese, dative-marking is typically reserved for beneficiary arguments only.

9.5.2 Case-marking in possessive constructions

The A argument in a DIP possessive construction is often not case-marked (cf. section 7.5.2), but we have already seen that a dative-marked argument can also be interpreted as a possessor - cf. e.g. (346c). The competition between the constructions A-V-P and DAT-A-V-P is not trivial. One important element here is that the transitive possessor verb *te* and the intransitive existential *te* are the same in DIP. Consider the following examples:

(347) a. *yo te ũ irmã i doy irmãw Go.*
1s *te* one sister and two brother Goa
'I have one sister and two brothers in Goa.'

b. *a mĩ te ũ irmã i doy irmãw Go.*
DAT 1s.OBL *te* one sister and two brother Goa
'I have one sister and two brothers in Goa.'

It seems clear that (347a) fills a monotransitive template. On the other hand, if one admits that the Dative-marked element in (347b) is an adjunct, then we could interpret

the clause as intransitive, *ũ irmã i doy irmãw* as S and the verbal form as an occurrence of existential *te*. Whatever the interpretation, this type of construction is quite typical of South Asian languages,[23] and Gujarati is no exception. Consider examples (345a,b) above, as well as (348) - adapted from Cardona (1965:96):

(348) *e   maṇes-ne be   chokr-a ch-e.*
       DET man-DAT two child-m EXS-3s
       'That man has two sons.' [Lit. 'To that man there are two sons.']

Dative-marked possessors are usually reserved for the expression of *natural possession* (mental objects and relatives). In Gujarati - and in various South Asian languages (see e.g. Mahajan 2004) - the possessor may also be locative-marked. The postposition *pase/paase* 'near' is used to indicate the temporary possession of physical objects (349a) and the locative suffix *-mã/mãã* for inbred virtues and vices (349b) - examples provided in Mistry (2004:6):

(349)   a. [*raaj paase*] *paisaa / be   gaaDi ch-e.*
             R.   near   money   two car   EXS-3s
             'Raj has money/two cars.'

         b. [*raaj-mãã*] *buddhi   / adekhaai ch-e.*
             R.-LOC   intelligence jealousy EXS-3s
             'Raj has intelligence/jealousy.'

In DIP, the A argument in comparable possessive constructions can be marked with one of the comitative prepositions *ku* or *jũt də* - see 7.5 for examples. At first sight, the fact that comitative case is selected in DIP would appear to cancel the parallelism with Gujarati, but consider that *jũt də* is also a locative preposition semantically very close to Guj. *pase/paase* 'near' (see 9.5.3). It is likely that *jũt də* was originally selected for this function on account of its locative semantics. Given the additional function of *jũt də* as a comitative marker, its comitative equivalent *ku* was at a later stage also allowed in this function. As case-markers of possessors, there is in fact a frequency differential which favours *jũt də* over *ku*, which supports the developmental scenario proposed.

In Portuguese, possessors are never dative-marked or locative-marked. In this respect, then, DIP sides clearly with Gujarati.

---

[23]There is some debate concerning the subjecthood of dative-marked arguments of possessive constructions in South Asian languages. Jayaseelan (1990, 2004), for instance, approaches the issue of Dative 'subjects' in Malayalam possessive constructions and concludes that the Dative-marked element cannot be said to be the clause's subject. Other authors (see e.g. Mahajan 2004, Mistry 2004), on the other hand, interpret Dative-marked (or other) arguments of possessive constructions as grammatical subjects.

9.5.3 The issue of complex adpositions

Despite the obvious differences in placement, there is a striking parallelism between Portuguese complex prepositions and Gujarati complex postpositions. In Portuguese, these consist of a semantic element combined with the relational preposition *de* 'of' (or rather a *de*-PP): e.g. *antes de* 'before [Lit. 'before of']', *ao pé de* 'near [Lit. 'at the foot of']', *em cima de* 'on top of', *em vez de* 'instead of [Lit. 'in turn of']'. In Gujarati, the relational element is the nominal affix *-n-*, which must be followed by gender and number suffixes (see Cardona and Suthar 2003:677) and which serves to link a given referent to various postposed elements, resulting in complex postpositions (Mistry 2004:2): e.g. *-n-i sathe* 'together with' - which functions as a comitative marker -, *-n-i pase* 'near, by' (see 9.5.2), *-n-e bədle* 'instead of', *-n-i upər* 'on top of'. In both Portuguese and Gujarati, the element identified as a generic relator is used as a genitive marker. In Gujarati, when the referent involved in one of these complex postpositional templates is pronominal, the pronoun takes a form similar to that of possessives - see Table 9.5 in section 9.3.1.

From a formal point of view, the DIP complex prepositions (for which, see 6.7 and 7.4) are clearly derived from Portuguese, not only because they are prepositional rather than adpositional but also because the relational element is the preposition *də*: e.g. *ãt də* 'before', *pərt də* 'near (of)', *sim də* 'on top of', *vez də* 'instead of'. On the other hand, there are certain patterns of use and variation which can only be understood if one recognises a pragmatic correspondence between particular DIP complex prepositions and particular Gujarati complex adpositions. I will illustrate this by exploring the use of DIP *jũt də*.

DIP *jũt də* derives from the Portuguese complex preposition *junto de*, which is a locative preposition meaning 'near, next to'.[24] In DIP, however, *jũt də* combines the functions of a locative preposition ('near, next to') and that of a comitative marker ('together with') - see 7.4.8. It therefore conflates the functions of two separate complex postpositions in Gujarati: *-n-i pase* 'near, by' and *-n-i sathe* 'together with'. One of the characteristics of the latter postposition in Gujarati, corresponding to the comitative construction, is that while personal pronouns always occur in their genitive form (e.g. *mari sathe* 'with me'), nouns consistently shun the relational element (e.g. *Ramesh sathe* 'with Ramesh' / *FM radio sathe* 'with FM radio'). This may begin to explain the fact that, in DIP, expressions with comitative *jũt də* plus a nominal participant can easily omit the relational preposition *də*. In addition, *jũt də* features in most instantiations of a possibility in DIP to alter the prototypical order of elements in a complex adposition template: e.g. *jũt + də +* N > (*də*) + N + *jũt*, therefore approaching its Gujarati equivalent not only in function but also in placement relative to the noun - cf. 7.4.8 and 7.4.10. In truth, other complex prepositions allow this type of inversion and/or *də* omission (see 7.4.10), but none as freely as *jũt də*. Crucially, other complex postpositions in Gujarati also admit the absence of case-marking on nouns but not on pronouns (e.g. *mari upper* 'on me', but *Ramesh ni upper* / *Ramesh upper* 'on Ramesh'); absence of the relational element in nouns is however optional for all these postpositions and only obligatory in the case of the comitative *sathe*.

---

[24]Ptg. *junto de* is not really used to express comitative, but the element *junto(s)/junta(s)* does mean 'together' (e.g. *nós estudámos juntos* 'we studied together'). The standard comitative marker in Portuguese is *com* 'with', but it is not uncommon for *junto* and *com* to be combined into a novel (emphatic) complex comitative preposition: *junto com* 'together with'.

9.6 Aspects of syntax

This section explores some aspects of DIP syntax in search of parallels with either Portuguese or Gujarati. The issue is much too vast to be covered here, and therefore this comparative analysis is simply illustrative. The various subsections deal with the syntax of content questions (9.6.1), copular constructions (9.6.2), comparative constructions (9.6.3) and the issue of the passive (9.6.4). The parallelism between Gujarati and DIP conjunct verbs is explained in 9.6.5, and the origins of DIP negative concord are explored in 9.6.6. Finally, 9.6.7 analyses the form of utterance complements.

9.6.1 Content question formation

Portuguese is somewhat flexible regarding the placement of the interrogative element (be it an interrogative pro-form or an interrogated NP) in content questions. The sentences in (350) illustrate Portuguese interrogatives indicating the optional positions of the interrogative elements (given in square brackets):

(350) a. [*quant-a-s   vez-es*] *já   *[*quant-a-s   vez-es*] *te avis-ei*
           how.many-f-p time-p already how.many-f-p time-p 2s warn-PST.1s
           [*quant-a-s   vez-es*]*?*
           how.many-f-p time-p
           'How many times have I already warned you?'

      b. [*como é   que*] *tu vi-este   *[*como*]*?*
          how   COP.PRS.3s CMP 2s come-PST.2s how
          'How did you come?'

Despite the variation observed, the prototypical position is initial (Mateus et al. 2003:464*ff*). In Gujarati as well as most Indo-Aryan languages (see Masica 1991:386*ff*), on the other hand, the interrogative elements are more strictly bound to one position, which is the focus position immediately preceding the VP.[25] The examples in (351) were taken from Cardona (1965:105;142):

(351) a. *təhm-e mumbəi-maN kyahre hə-š-o?*
           2s-AG M.-LOC   when   EXS-FUT-m.s
           'When will you be in Bombay?'

      b. *rameš šuN   kam kəre ch-e?*
          R.   what work do   AUX.PRS-3s
          'What work does Ramesh do?'

---

[25] The VP is in this case meant to include close-knit VP modifiers, such as clausal negators

    c. *təhm-e kem aw-y-a?*
       2p-AG how come-PST-m.p
       'How did you come?'

On rare occasions, an argument may retain its hold onto the focus position, in which case the interrogative element will precede it. Consider example (352), from Lambert (1971:120), quoted in Masica (1991:387):

(352)   *āp kyāre ghēr javānā ch-o?*
       2s when house come   AUX-2p
       'When are you coming home?'

With respect to the placement of interrogative elements in content questions, DIP mirrors Gujarati very closely. As described in sections 7.2.4 and 7.6.2.3, the DIP interrogative element is bound to the pre-verbal focus position. As an illustration, consider example (353), the translation equivalent of the Portuguese question in (350b) and its Gujarati counterpart in (351c):

(353)   *use kom vey?*
       2s   how come.PST
       'How did you come?'

The order of elements in this question cannot be modified, unless the displaced constituents are bracketed off by intonational breaks. It is interesting to notice that, even though DIP aligns with Portuguese with respect to the basic word order of simple declaratives (cf. 9.4), the language clearly sides with Gujarati syntax when it comes to content question formation.

9.6.2   Copular constructions

In Gujarati non-verbal predication, the copular verb may be omitted if the polarity of the clause is positive. This is exemplified in (354):

(354)   *Subodh-n-i nəv-i kavita ghəṇ-i sundər   (ch-e).*
       S.-of-f   new-f poem much-f beautiful COP.PRS-3p
       'Subodh's new poem is very beautiful'

The copula is required in negative polarity contexts. The copular auxiliary has specific negative forms (e.g. *nəthi*) which are used in such cases.

As far as Portuguese is concerned, a copula is always required to accompany non-verbal predication. This language makes use of an individual-level copula (the verb *ser*)

and a stage-level copula (the verb *estar*), with somewhat different semantics. The former indicates that the truth-value of the predication is perceived as stable, whereas the latter constructs a predication believed to be temporary.

DIP sides with Portuguese almost entirely with respect to the expression of non-verbal predication. The copula is a verbal element which cannot be omitted, and the language also makes use of two different copulas (*ɛ/ɛr* and *te/tiŋ*, see section 7.1.3) indicating stable and temporary conditions respectively. There is an important aspect concerning the DIP stage-level copula, however, which does show Gujarati influence. In Gujarati, the verbal form which functions as a copula is also an auxiliary verb, and it is also used in existential, locative and possessive constructions. The DIP form of the stage-level copula *te/tiŋ* appears not to be immediately derived from the Portuguese copula *estar* or any of its inflected forms but from the possessive verb *ter* 'to have' [PST *tinha* 'had']. In modern DIP, the functional scope of *te/tiŋ* coincides with that of its Gujarati counterpart in that it is not only a copular verb and a possessive verb but also an existential, an auxiliary and a locative verb.[26]

9.6.3 Comparative construction

The comparative construction is Portuguese makes use of a comparative particle (*do*) *que*, which precedes the standard of comparison. The order of elements is very much fixed: the comparee is followed by the copular verb, followed by the parameter (modified by either *mais* 'more' or *menos* 'less'), the comparative particle and finally the standard of comparison. An example is given in (355):

(355) *aquel-e   é   mais barat-o   (do) que est-e.*
       DEM$_d$-m COP.PRS.3s more cheap-m COMP   DEM$_p$-m
       'That one is cheaper than this one.'

In Gujarati, comparison always manifests superiority and involves a locational strategy in that the standard is expressed in ablative case, marked with the suffix *-thi*; alternatively, a postposition *kertã*, exclusive to comparative constructions, can be used. Both the standard of comparison and the comparee precede the parameter of comparison, but their order is interchangeable. Notice the following examples (from Cardona 1965:91), in which the standard of comparison is enclosed in square brackets:

(356) a. *pel-ũ   [a-thi]   səst-ũ   ch-e.*
           DEM-N DEM-ABL cheap-N COP.PRS-3p
           'That one is cheaper than this one.'

---

[26]Clements and Koontz-Garboden (2002:200) report the same pattern of macrofunctionality for Korlai IP, as well as Daman IP (with the exception of the copula), reflecting the strategies of Marathi and Gujarati respectively.

b. [*pel-a*   *ghər kertã*] *a*   *ghər sundər*   *che.*
DEM-OBL house *kertã* DEM house beautiful PRS-3p
'This house is more beautiful than that house.'

In (356a), the ablative-marked standard of comparison (in this case a pronominal argument) occurs after the comparee, while in (356b) it precedes the comparee. The relative position of the two arguments is always flexible, not a function of the two standard-marking strategies. Notice also that there is no modification whatsoever of the expressed parameter.

The DIP construction is, in a sense, a mix of both the Portuguese and the Gujarati systems - see section 7.11.1. The DIP comparative involves a particle construction (using the comparative particle *ki*) with a modified parameter (modified by *may*, very rarely by *men*) and an ablative-marked standard (marked by the preposition *də*). The order of constituents is close to that of Portuguese. Contrast the DIP example in (357) with the equivalent Portuguese and Gujarati clauses given in (355) and (356a) respectively:

(357)   *ikə*   *ε*   *may barat ki*   *d-es.*
       DEM_d COP_i.NPST more cheap COMP ABL-DEM_p
       'That one is cheaper than this one.'

9.6.4   Passive construction

Gujarati has both a morphological passive and a periphrastic passive. As explained in section 9.2.2, the verbal passive affix is *-a-*, appended to the root of the verb (implying root alterations if they are vowel-final or contain the segment /a/) - see example (358a). The periphrastic passive involves a participle form of the verb and a form of the verb *avũ* 'to come' providing TAM information - this is exemplified in (358b), adapted from Cardona and Suthar (2003):

(358)   a.   *mara-thi a*   *kam nəhi kər-a-y.*
           1s-ABL   DEM work NEG do-PASS-3p
           'I won't be able to do this job.' [Lit. 'This job won't be done by me']

   b.   *mara-thi a*   *kam nəhi kər-v-a-mã*   *av-e.*
           1s-ABL   DEM work NEG do-PTCP-m-LOC come-3p
           'I won't be able to do this job.' [Lit. 'This job won't be done by me']

The agent argument is modified by the ablative-marker *-thi* but no particular case-marking is required on the undergoer. The Gujarati passive does not involve any particular alteration to the prototypical word order either. In this language, a passive construction is often used simply to express (in)ability, as in the previous examples.

In Portuguese, the passive construction is strictly periphrastic, involving a participle and a form of the verb *ser* 'to be' inflected for TAM. The sentences in (359) exemplify this

construction:

(359)   a.   *tu já   fo-ste   avisa-d-o.*
             2s already be-PST.2s warn-PTCP-m
             'You have already been warned.'

        b.   *a   pedra foi   atira-d-a   por el-e.*
             ART.f stone be.PST.3s throw-PTCP-f by 3s-m
             'The stone was thrown by him.'

The Portuguese passive is an effective strategy to demote the agent of the predicate and enhance the pragmatic centrality of the undergoer. The undergoer appears in subject (i.e., first) position; the agent need not be expressed at all, but whenever expressed it is embedded in a PP introduced by *por* 'by'.

In DIP, the most frequent strategy to demote a particular argument which is deemed non-essential for the speech-act is simply to omit it - see section 7.8.2. DIP lacks any formal means to passivise - see chapter 7 fn. 14. The absence of a passive construction is therefore one of the features with respect to which DIP departs from both Gujarati and Portuguese.

9.6.5   Conjunct verbs

Section 7.1.4 described the DIP constructs called *conjunct verbs*, i.e. complex verbs which result from the combination of a light verb (*faze* 'to make', in the case of DIP) and a non-verbal element. These constructs function as a single predicative unit: the non-verbal element does not occupy an argument slot, and the combination often acquires very specific semantics. An equivalent strategy is very widespread among modern Indo-Aryan languages.[27] In Gujarati, the most common light verb involved is *karvũ* 'to make' (equivalent to DIP *faze*). Examples include *fon karvũ* 'to make a phone call' [Lit. 'to make telephone'], *yad karvũ* 'to remind' [Lit. 'to make memory'] and *saf karvũ* 'to clean' [Lit. 'to make clean']. Consider the following example, from Cardona and Suthar (2003:688):

(360)   *mẽ   nədi   par kər-i.*
        1s.AG river[f] up make-PFV.f
        'I crossed the river.'

---

[27]A different label for these complexes in the literature on South Asian languages is *nominal-verb complexes*. These should however not be mistaken for *compound verbs*, which refer to the combination of two verb forms, even though the distinction is not clear in the work of some scholars, see for instance Koul (2003) concerning Kashmiri or Kachru (2008) on Hindustani. Masica (1991:368*ff*) provides an overall account of the phenomenon across the Indo-Aryan language family. For Gujarati, see Cardona (1965) and Cardona and Suthar (2003). Kachru (1982) deals with conjunct verbs in both Hindi-Urdu and Persian, Pandharipande (2003) reports them in Marathi, and Shackle (2003) in Panjabi.

In (360), there is evidence that *par kər-* 'up make' functions as a unit, given that the verb agrees in gender and number with *nədi* 'river'. The fact that Portuguese does not feature a parallel possibility means that DIP aligns with Gujarati with respect to the use and productivity of conjunct verbs.

9.6.6   Negative concord

Unlike DIP, Gujarati does not allow double negation. The occurrence of a negated NP does away with any additional clausal negator.

In Portuguese, on the other hand, negated NPs (e.g. *nenhum aluno* 'no pupil') or negative pronominals (e.g. *ninguém* 'nobody') co-occur with clausal negation - therefore producing negative concord - but only in the event that the negated constituent appears after the clausal negator. This is exemplified in (361):

(361)　　a.　*ninguém / nenh-um alun-o　aparece-u.*
　　　　　　　nobody　　NEG$_{ct}$-m pupil-m show.up-PST.3s
　　　　　　　'Nobody / no pupil showed up.'

　　　　　b.　*não　aparece-u　　ninguém / nenh-um alun-o.*
　　　　　　　NEG$_{cl}$ show.up-PST.3s nobody　　NEG$_{ct}$-m pupil-m
　　　　　　　'Nobody / no pupil showed up.'

In (361a), the negated NP/negative pronoun occurs clause-initially and, as such, no clausal negation is licensed. In (361b), on the other hand, the fact that the verb precedes the negated constituent triggers double negation; in these circumstances, double negation is not only possible but obligatory.

One may be inclined to attribute DIP negative concord to its main lexifier language, but the truth is DIP differs from Portuguese (and from Gujarati) in that negative concord is obligatory whenever a negated constituent is realised, irrespective of its position relative to the clausal negator - see section 7.7 for further details.

9.6.7   Utterance complements

This section deals solely with a subtype of complement clauses, viz. utterance complements, those which report an utterance as complements of verbs of saying such as 'to tell', 'to say' or 'to declare'. In DIP, the most common utterance verbs are *fala* 'to say' and *dize* 'to say'. Their clausal complements are expressed in direct speech (with rare exceptions) and may or may not be introduced by the complementiser *ki*. Like other complement clauses, the complements of utterance verbs normally follow the main clause - see section 7.10.1 for a full description of these constructions.

In Portuguese, complements of utterance predicates may occur in direct or indirect speech, but the former never admit the complementiser *que* while the latter require its

presence:

(362) a. *o*   *João di-sse:*   *"espera   por mim".*
           ART.m J.   say-PST.3s wait.IMP for 1s.OBL
           'João said: "wait for me".'

       b. *o*   *João di-sse*   *que   esper-ass-e*   *por ele.*
           ART.m J.   say-PST.3s CMP wait-SBJV-IPFV.1s for 3s
           'João told me to wait for him.' [Lit. 'João said that I (should) wait for him.']

The sentence in (362a) features an utterance complement in direct speech, while in (362b) the complement is introduced by the complementiser *que* and appears in indirect speech. Whereas the order of elements is the same in Portuguese and DIP (i.e., the utterance complement follows the main clause and the eventual complementiser is clause-initial), the two diverge significantly in that DIP does not select indirect speech and allows the co-occurrence of direct speech and the complementiser.

The equivalent strategy in Gujarati sheds some light onto the DIP/Portuguese divide. In Gujarati, complement clauses in general can be placed either prior to the main clause or after it, the latter being the most common placement. The subordinators used in each of the constructions are different. The element *ke*[28] is clause-initial and introduces complements postposed to the main clause, while *ɛm* is clause-final and follows a complement clause preposed to the main clause. These competing possibilities are exemplified in (363) - adapted from Cardona and Suthar (2003:690):

(363) a. *kišor-e kəh-y-ũ*   [*ke   mar-i   rah jo-j-o*].
           K.-AG say.PFV-n CMP POSS.1s-f way look-IMP-p
           'Kishore told me to wait for him.' [Lit. 'Kishore said that "look my way".']

       b. [*mar-i   rah jo-j-o   ɛm*] *kišor-e kəh-y-ũ.*
           POSS.1s-f way look-IMP-p CMP K.-AG say.PFV-n
           'Kishore told me to wait for him.' [Lit. '"Look my way" Kishore said.']

The complements in sentences in (363) - the sequences in square brackets - are utterance complements and appear in direct speech form whatever their position relative to the main clause (see also Masica 1991:403). Indirect speech utterance complements are in fact also available to Gujarati speakers, although they are dispreferred to direct speech.

The structural possibilities of utterance complements in DIP therefore show influence from both Gujarati and Portuguese: the fact that the complementiser is optional owes to Portuguese while the fact that the complement is consistently expressed in direct speech (even in combination with a complementiser) appears to be modelled on Gujarati.

---

[28] *Ke* also occurs with different functions, viz. as a disjunctive conjunction meaning 'or' and, in colloquial speech, as a tag (which is probably a reduction of *ke nahi* 'or not').

9.7 Concluding remarks

Any language which operates and evolves in a multilingual space is exposed to a highly diverse pool of available features and, particularly in the relative absence of standardising pressure, more or less free to select features and recombine them into new systems. The challenge for contact linguistics is therefore to formulate principles and criteria to account for the actual selections made by particular languages. This subsection intends to ascertain to what extent the comparison effected in the present chapter may contribute to our understanding of the mechanisms of language change. As it turns out, the evidence unearthed for DIP is more iconoclastic than conclusive. Let us begin by determining whether it is possible to recognise general guiding principles behind the complex typological allegiances of DIP.

9.7.1 The typological allegiances of DIP

The most striking observation concerning the comparative study effected here is that, in typological terms, DIP sides with Gujarati, or Portuguese, or departs from both in ways that are anything but straightforward. It is not possible to quantify the contribution of each of the ancestor languages to the typological make-up of DIP, nor was that the intention of this study, but it is possible to try and understand whether or not the contributions of one or the other are restricted to different domains.

In some cases, one can abstract general trends shaping the alignment of DIP with Portuguese or Gujarati. The most powerful generalisation is the one that attributes to Portuguese the role of privileged lexifier. As we have seen before, for all the ease with which DIP-speakers juggle lexical material from various sources, Portuguese has provided the vast majority of the most current lexical labels in DIP and, significantly, the forms of nearly all of its functional items. Another clear dichotomy operates within the domains of phonology and phonetics. As noticed in 9.1.3.5, in general the phonological system of DIP owes much to Classical Portuguese, though not without various systematic alterations which, however, cannot clearly be assigned to Gujarati influence. The contribution of Gujarati in this domain is vast but mostly confined to the realm of phonetic realisation.

In other domains, such as syntax, a similar division of tasks is elusive. Recall, for instance, that while DIP sides with Portuguese with regard to its basic word order (section 9.4), it follows the syntactic precepts of Gujarati very closely with respect to question formation (section 9.6.1).

It is also often the case that DIP mirrors a particular structure or possibility from one of its ancestor languages but introduces alterations. One case in point is that of negative concord which, despite being shared with Portuguese, has in fact a degree of acceptance (and obligatoriness) in DIP which goes beyond the possibilities of double negation in Portuguese (see section 9.6.6).

It is interesting to notice that certain characteristics of DIP combine elements from both Gujarati and Portuguese. Recall for instance the structure of the comparative in DIP, which makes use of both a Portuguese-style comparative marker plus parameter modification, and Gujarati-style ablative case on the standard of comparison (section 9.6.3); or the variation observed in the expression of utterance complements (section 9.6.7). In addition, while

shunning the S-O-V word order of Gujarati, DIP retains a very clear preverbal focus position modelled on this language, giving rise for instance to innovative constructions such as those involving constituent doubling (section 7.8.3).

Finally, DIP often differs from both Portuguese and Gujarati, as for instance with regard to stress assignment rules (section 9.1.3.4), the absence of a passive construction (section 9.6.4) or the relative marginality of nominal/verbal morphology (section 9.2). These singularities beg the question whether they are DIP innovations or whether they can be accounted for with reference to those strands of initial input (e.g. the putative Portuguese-based pidgin, other Indian or European languages, non-standard varieties of all the intervening languages) which are less well defined than Gujarati or Portuguese (see section 3.3). Many of these questions cannot be answered at this stage, see 9.7.2 below.

The realisation of the typological 'hybridity' of DIP poses a serious challenge to almost any 'one-size-fits-all' theory of contact language formation. Traditional views of high-contact languages as consisting of the combination of one language's lexicon with another language's grammar[29] cannot fully account for the panorama in DIP; we have established that, alongside Indo-Aryan syntactic structures, semantic patterns, pragmatic strategies and so forth, DIP also sides strongly with Portuguese in domains other than the lexicon. The truth is the overall picture begs the recognition that high-contact varieties may be more intricately mixed than often assumed.

9.7.2 Diachrony

The comparative study carried out in this chapter (as well as the description of DIP in general) highlighted some important diachronic facts. One of the most relevant is the fact that the linguistic input involved in the initial stages of language contact (the early/mid-16th century, see chapter 3) has had an impact on DIP which, in various domains, has not been supplanted by later change affecting the linguistic landscape of Diu. This was made particularly evident with the discussion of DIP phonology, which revealed that the language retains certain segments and oppositions which can only be traced back to Classical Portuguese, not to Modern Portuguese (see section 9.1) or Gujarati. A similarly point was made in the discussion of DIP lexicon, in particular with regard to the retention of Portuguese archaisms (see section 8.7.1).

It is important to understand, with regard to the diachrony of DIP, that the languages most centrally involved in its formation (Gujarati and Portuguese) have remained in constant contact with DIP to this day. Therefore, the fact that the influence of the earliest stages of these languages is still visible reminds us that the moment of initial contact is often particularly determinant: the *founder principle* of evolutionary biology and first applied to the study of language contact by Mufwene (1996, 2001).

This certainly does not mean DIP crystallised at any point of its history. Despite the absence of early sources, it is possible to observe change even just when we confront

---

[29]An instantiation of this view enforms the theory that language contact primarily entails relexification (i.e. the replacement of a language's structures and semantics with another language's lexical labels), as defined for instance in Lefebvre (1998, 2004). See DeGraff (2002) for a survey and critique of the concept of relexification and its application to creole genesis.

the present state of the language with the 19th-century corpus available (see section 1.3). Section 7.2.7.1 described how the standard pluralising strategy in the 19th century involved nominal reduplication, which has almost entirely disappeared from modern DIP. Further comparison will reveal additional evidence of recent change, including the loss of a preverbal Past tense marker *jə* (obligatorily preposed to every inflected Past perfective verb forms, in the 19th c.) or the alteration of certain possessive and pronominal forms (see this chapter fn. 16).

In general, though, the absence of early records of the language makes it impossible to ascertain whether certain traits of DIP were selected from the original linguistic cauldron, established in the earliest stages of language contact, or whether they result from subsequent processes of change. As a result, one cannot be precise regarding such issues as the 'primary' source of linguistic material for DIP or the 'directionality' of subsequent change. Nonetheless, the sociohistorical data at our disposal (see chapter 3) and recent observations suggest that an important constant throughout the history of DIP has been the tension between the normative pull of standard Portuguese and the tendency towards metatypy (to use the terminology proposed by Ross 1996, 2001) or convergence (in the sense of e.g. Gumperz and Wilson 1971) with Gujarati.

Multilingualism has been the norm among native speakers of DIP, from the very inception of the language to the present day, and multilingualism is a precondition for the operation of linguistic convergence (see Sridhar 2008). The crucial factor here is that, from very early on, the DIP-speaking community has been mostly made up of local Diuese for whom Gujarati was a strong second language (see 3.2.3.1) and Portuguese was not.[30] Therefore, for most of the history of DIP, Portuguese has played the role of a prestigious norm proficiently spoken by either non-speakers of DIP or a small minority of native-speakers, while Gujarati has been a highly fluent second-language for the vast majority (if not the whole) of the DIP-speaking population. Their influence on DIP (beyond the formative stages) therefore operated through rather different mechanisms. Gujarati exerts a converging pull naturally arising from its widespread use within the DIP-speaking community; the function of Portuguese as a norm, on the other hand, is somewhat weakened by the fact that the language has had only limited currency in Diu, in the recent past at least.

The result of the conflict between the normative and converging forces is to 'arrest' the expected process of metatypy towards Gujarati, which may explain some of the seemingly incongruous patterns of typological allegiance described in section 9.7.1. The decline in the presence and role of Portuguese in Diu which has taken place since 1961 (see chapter 2) is likely to eventually tip the scales in favour of convergence with Gujarati, but the relevant variables are entirely social and, as such, dependent on such unpredictable factors as sociopolitical feeling, educational efforts, population movement or affection.

---

[30] I do not mean to imply, as some scholars have concerning the formation of creole languages, that the motivation behind the formation of DIP was some sort of 'missed' attempt at acquiring Portuguese - see Baker (1990) for an assessment of this notion. Much on the contrary, it is fair to recognise that DIP is the linguistic manifestation of a community which has been highly succesful as mediator between rather distinct cultural worlds - i.e., a community negotiating 'multiple identity alignment', in the formulation of Ansaldo (forthcoming).

9.7.3 Further study

The data provided in the present description opens up a number of different avenues for further research. One of the most promising involves a systematic comparison with other contact varieties which have arisen in similar typological and/or social circumstances. Naturally, the most obvious objects of comparison are the other IP varieties and the Portuguese-lexified creoles of Asia, in general. One cannot expect to find patterns of feature selection and adaptation being replicated in all languages, but a comparative study of this nature will permit the exploration of a number of different questions:

a) the exact features available to the creators of these languages at a given point in time, given that the feature pool which fed into the different varieties was probably overlapping to some extent;

b) the *directionality* (if any) of the spread of Portuguese-lexified contact varieties across Asia, associated to the intuitive hypothesis that, along the eastward political expansion of the Portuguese in the early 16th century, population and linguistic transfer must have also proceeded eastwards;

c) the extent to which the Portuguese-lexified contact varieties of Asia developed in isolation, and to what extent they influenced each other; recall the 'recíproca transfusão parcial' [partial reciprocal transfusion] hypothesis advanced in Dalgado (1917) - see chapter 3 fn. 39;

d) the hypothesis that a high degree of overlap in the relevant typological contexts resulted in a high degree of similarity between the different varieties.

It would be advantageous to bring other contact varieties of (South) Asia on board such a contrastive study, such as e.g. Bazaar Hindi, Sri Lanka Malay or Nagamese (see Smith 2008a). It would be particularly interesting to determine how much of the specifity of DIP can be assigned to areal factors. As an example, notice that even a superficial comparison between DIP and Sri Lanka Malay reveals striking similarities in terms of the adoption of Indic case-marking strategies (cf. Aboh and Ansaldo 2007, Ansaldo 2008). Given that so much of the outcome of language contact is determined by social factors, I am convinced that further analysis of the social contexts of (South) Asian contact varieties will clarify the mechanisms by which these languages formed and developed.

Bibliography

Abbi, Anvita. 2008. Tribal languages. In *Language in South Asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 153–174. Cambridge: Cambridge University Press.

Aboh, Enoch O., and Umberto Ansaldo. 2007. The role of typology in language creation. In *Deconstructing Creole*, ed. Umberto Ansaldo, Stephen Matthews, and Lisa Lim, 39–66. Amsterdam: John Benjamins.

Aikhenvald, Alexandra, and Robert Malcolm Ward Dixon. 2006. *Serial Verb Constructions. a cross-linguistic typology*. Oxford: Oxford University Press.

Albernaz I, João Teixeira. c. 1648. Plantas das Cidades e Fortalezas da Conquista da India Oriental. Codex: Bayerische Staatsbibliothek, München, Cod. Mon. Icon. 162.

Albuquerque, Afonso de. 1774. *Comentários do grande Afonso Dalboquerque capitão geral que foi das Índias Orientaes em tempo do muito poderoso Rey D. Manuel, o primeiro deste nome*. Lisbon: Regia Officina Typographica.

Altekar, A. S. 1924-1925. A history of important ancient towns and cities of Gujarat and Kathiawad from earliest times to about 1300 A. D. *Indian Antiquary* LII-LIV.

Anderson, Gregory. 2006. *Auxiliary verb constructions*. Oxford: Oxford University Press.

Andrade, Francisco de. 1589. *O primeiro cerco que os turcos puserão ha fortaleza de Diu nas partes da Índia*. Coimbra.

Annamalai, E. 2003. The opportunity and challenge of language documentation in India. In *Language documentation and description*, ed. Peter K. Austin, volume 1, 159–167. London: SOAS.

Ansaldo, Umberto. 2008. Sri Lanka Malay revisited: genesis and classification. In *Lessons from documented endangered languages*, ed. K. David Harrison, David S. Rood, and Arienne Dwyer, 13–42. Amsterdam: Benjamins.

Ansaldo, Umberto. forthcoming. *Contact languages. Ecology and evolution in Asia*. Cambridge University Press.

Ansaldo, Umberto, Jan Don, and Roland Pfau. 2008. Parts of Speech: Particulars, universals and theoretical constructs. *Studies in Language* 32:505–508.

Ansaldo, Umberto, Stephen Matthews, and Lisa Lim, ed. 2007. *Deconstructing Creole*. Amsterdam: John Benjamins.

Araújo, Maria Benedita de Almeida. 2008. *Campanhas da Índia. Sofala, Goa e Malaca*. Matosinhos: Quidnovi.

Arends, Jacques, Pieter Muysken, and Norval Smith, ed. 1995. *Pidgins and Creoles - an introduction*. Amsterdam: John Benjamins.

Arends, Jacques, and Matthias Perl, ed. 1995. *Early Suriname Creole texts; A collection*

of 18th-century Sranan and Saramaccan documents. Frankfurt: Vervuert.

van der Auwera, Johan. 1998a. Introduction. In *Adverbial constructions in the languages of Europe*, ed. Johan van der Auwera and Dónall Ó Baoill, 1–24. Berlin: Mouton de Gruyter.

van der Auwera, Johan. 1998b. Phasal adverbials in the languages of Europe. In *Adverbial constructions in the languages of Europe*, ed. Johan van der Auwera and Dónall Ó Baoill, 25–145. Berlin: Mouton de Gruyter.

Baker, Philip. 1990. Off target? *Journal of Pidgin and Creole Languages* 5:107–119.

Bakker, Peter. 2002. Some future challenges to Pidgin and Creole studies. In *Pidgin and Creole studies in the twenty-first century*, ed. Glenn Gilbert, 69–92. New York: Peter Lang.

Baptista, Marlyse. 2005. New directions in pidgin and creole studies. *Annual Review of Anthropology* 34:33–42.

Barbosa, Jorge Morais. 1965. *Études de phonologie Portugaise*. Lisbon: Junta de Investigações do Ultramar.

Barros, João de. 1615. *Quarta década da Ásia - Dos feitos que os Portuguese fizerão no descobrimento, e conquista dos mares, e terras do Oriente*.

Barros, Teresa Leitão de. 1954. *As heroínas de Diu*. Lisbon: S.N.I.

Batalha, Graciete. 1959. *Estado actual do dialecto macaense*. Coimbra: Universidade de Coimbra.

Baxter, Alan. 1983. Creole universals and Kristang (Malacca Creole Portuguese). In *Papers in pidgins and creole linguistics*, number 3 in Pacific Linguistics, 143–160.

Baxter, Alan N. 1988. *A Grammar of Kristang (Malacca Creole Portuguese)*. Canberra: Pacific Linguistics.

Baxter, Alan N. 1990. Notes on the Portuguese Creole of Bidau, East Timor. *Journal of Pidgin and Creole Languages* 5:1–38.

van den Berg, Margot, and Jacques Arends. 2004. Court records as a source of authentic early Sranan. In *Creoles, contact and language change: Linguistic and social implications*, ed. Geneviève Escure and Armin Schwegler. Amsterdam: John Benjamins.

Berrenger. 1811. *A grammatical arrangement on the method of learning the corrupted Portuguese as spoken in India*. Government Press, 2nd ed. edition.

Bhat, D. N. S., Regina Pustet, and Boulder. 2000. Adjective. In *Morphologie, Ein internationales Handbuch zur Flexion und Wortbildung*, ed. Geert Booij, Christian Lehmann, and Joachim Mugdan, 757–769. Berlin: Mouton de Gruyter.

Bhat, Darbhe Narayana Shankara. 2004. *Pronouns*. Oxford University Press.

Bhatt, Rakesh M., and Ahmar Mahboob. 2008. Minority languages and their status. In *Language in South Asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 132–152. Cambridge: Cambridge University Press.

Biker, Julio Firmino Judice. 1881. *Collecção de Tratados e concertos de pazes que o Estado da Índia Portugueza fez com os Reis e Senhores com quem teve relações nas partes da Ásia e Africa Oriental desde o principio da conquista até ao fim do século XVIII*. Lisbon: Imprensa Nacional.

Bocarro, António. 1635. Livro das plantas de todas as fortelezas, cidades e povoações do

Estado da India Oriental. Codex: Biblioteca Pública e Arquivo Distrital de Évora, COD. CXV/2-X.

Bolinger, Dwight L. 1958. A theory of pitch accent in English. *Word* 14:109–49.

Bourdieu, Pierre. 1982. *Ce que parler veut dire: l'économie des échanges linguistiques*. Paris: Fayard.

Bourdieu, Pierre. 1991. *Language and symbolic power*. Cambridge, Mass.: Harvard University Press.

Boxer, Charles R. 1961. *Four centuries of Portuguese expansion, 1415-1825: a succint survey*. Johannesburg: Witwatersrand University Press.

Boxer, Charles Ralph. 1963. *Race relations in the Portuguese colonial empire, 1415-1825*. Oxford: Clarendon Press.

Boxer, Charles Ralph. 1984. The *carreira da Índia* (ships, men, cargoes, voyages). In *From Lisbon to Goa, 1500-1750: studies in Portuguese maritime enterprise*, ed. Charles Ralph Boxer. London: Variorum Reprints.

Bragança Pereira, A. B. de. 1938. *Os portugueses em Diu*. Bastorá: Tipografia Rangel. Separata *O Oriente Português*.

Bragança Pereira, A. B. de. 1940. *Etnografia da India Portuguesa*, volume 2. Bastorá: Tipografia Rangel.

Brito, Raquel Soeiro de. 1998. *Goa e as Praças do Norte*. Lisbon: Comissão Nacional para as Comemorações dos Descobrimentos Portugueses.

Bybee, Joan, and Suzanne Fleischman. 1995. *Modality in grammar and discourse*. Amsterdam: Benjamins.

Cardeira, Esperança. 2005. *Entre o Português Antigo e o Português Clássico*. Lisbon: Imprensa Nacional - Casa da Moeda.

Cardona, George. 1965. *A Gujarati reference grammar*. Philadelphia: University of Pennsylvania Press.

Cardona, George, and Babu Suthar. 2003. Gujarati. In *The Indo-Aryan languages*, ed. George Cardona and Dhanesh Jain, 659–697. New York: Routledge.

Cardoso, Hugo C. 2003. Processes of change in the Portuguese-based lexicon of Saramaccan. Mphil thesis, Universiteit van Amsterdam.

Cardoso, Hugo C. 2006a. Challenges to Indo-Portuguese across India. In *Proceedings of the FEL X*, ed. R. Elangayan, R. McKenna Brown, Nicholas D.M. Ostler, and Mahendra K. Verma, 23–30. Mysore.

Cardoso, Hugo C. 2006b. Diu Indo-Portuguese at present. *Journal of Pidgin and Creole Languages* 21:359–370.

Castro, Ivo. 1991. *Curso de História da Língua Portuguesa*. Lisbon: Universidade Aberta.

Chauhan, R. R. S. 1995. *Africans in India: from slavery to royalty*. New Delhi: Asian Publication Services.

Clark, Eve. 1978. Locationals: existential, locative and possessive constructions. In *Universals of human language*, ed. Joseph H. Greenberg, volume 4, 85–126. Stanford: Stanford University Press.

Clements, Joseph Clancy. 1991. The Indo-Portuguese Creoles: languages in transition. *Hispania* 74:637–646.

Clements, Joseph Clancy. 1996. *The genesis of a language*. Amsterdam: John Benjamins.

Clements, Joseph Clancy. 2000. Evidência para a existência de um pidgin português asiático. In *Crioulos de base lexical portuguesa*, ed. Ernesto d'Andrade, Dulce Pereira, and Maria Antónia Mota. Braga: Associação Portuguesa de Linguística.

Clements, Joseph Clancy. forthcoming. Gradual vs. abrupt creolization and recent changes in Daman Creole Portuguese. In *Gradual Creolization: Studies celebrating Jacques Arends*, ed. Rachel Selbach, Hugo Cardoso, and Margot van den Berg. Amsterdam: John Benjamins.

Clements, Joseph Clancy, and Andrew J. Koontz-Garboden. 2002. Two Indo-Portuguese Creoles in contrast. *Journal of Pidgin and Creole Languages* 17:191–236.

Coelho, F. Adolpho. 1880. *Os dialectos românicos ou neo-latinos na África, Ásia e América*. Sociedade de Geographia de Lisboa.

Comrie, Bernard. 1976. *Aspect. an introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1978. Ergativity. In *Syntactic typology: studies in the phenomonology of language*, ed. Winfred P. Lehmann, 329–394. Austin: University of Texas Press.

Comrie, Bernard. 1989. *Language universals and linguistic typology*. Chicago: University of Chicago Press.

Corte Real, Jerónimo. 1574. *Sucesso do Segundo Cerco de Diu. Estando Dom Joham Mazcarenhas por capitam da fortaleza. Anno de 1546*. Lisob.

Couto, Diogo do. 1602. *Década Quarta da Ásia*. Republished in 1999.

Couto, Diogo do. c. 1571. Diálogo do Soldado Prático. Que trata dos enganos e desenganos da Índia.

Cristofaro, Sonia. 2005a. Purpose clauses. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 125. Oxford: Oxford University Press. Available online at http://wals.info/feature/125.

Cristofaro, Sonia. 2005b. Utterance complement clauses. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 128. Oxford: Oxford University Press. Available online at http://wals.info/feature/128.

Croft, William. 2000. Parts of speech as language universals and as language-particular categories. In *Approaches to the typology of word classes*, ed. Petra M. Vogel and Bernard Comrie, 65–102. Berlin: Mouton de Gruyter.

Croft, William. 2001. *Radical Construction Grammar*. Oxford: Oxford University Press.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press, 2nd edition.

Cuesta, Pilar Vázquez, and Maria Albertina Mendes da Luz. 1980. *Gramática da língua portuguesa*. Lisbon: Edições 70.

Cunha, Celso, and L. F. Lindley Cintra. 1984. *Nova Gramática do Português Contemporâneo*. Lisbon: Edições João Sá da Costa.

Östen Dahl. 1985. *Tense and Aspect systems*. Oxford: Basil Blackwell.

Dalgado, Sebastião Rodolfo. 1900. Dialecto indo-português de Goa. *Revista Lusitana* VI:63–84.

Dalgado, Sebastião Rodolfo. 1903. Dialecto indo-português de Damão. *Ta-ssi-yang-kuo*. Published in two *separatas*: 3:359-367 (1902); 4:515-523 (1903).

Dalgado, Sebastião Rodolfo. 1906. Dialecto indo-português do Norte. *Revista Lusitana* 9:142–166, 193–228.

Dalgado, Sebastião Rodolfo. 1913. *Influência do vocabulário português em línguas asiáticas*. Coimbra: Imprensa da Universidade.

Dalgado, Sebastião Rodolfo. 1917. Dialecto Indo-Português de Negapatão. *Revista Lusitana* 20:40–53.

Dalgado, Sebastião Rodolfo. 1921. Berço duma cantiga em Indo-Português. *Revista Lusitana* XXII. Separata.

Daniel, Michael, and Edith Moravcsik. 2005. The Associative Plural. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 36. Oxford: Oxford University Press. Available online at http://wals.info/feature/36.

Dasgupta, Probal. 2004. Some non-nominative subjects in Bangla. In *Non-nominative subjects*, ed. Peri Bhaskararao and Karumuri Venkata Subbarao, volume 1, 129–140. Amsterdam: Benjamins.

DeGraff, Michel. 2002. Relexification: a reevaluation. *Anthropological Linguistics* 44:321–414.

DeGraff, Michel. 2003. Against Creole exceptionalism. *Language* 79:391–410.

Delduque, Adelino. 1928. *Diu: breve notícia histórica e descritiva*. Lisbon: J. Rodrigues & C.a.

Dhruva, M. N. 1920. *Gujarati self-taught*. London: Marlborough. Revised edition by T. N. Dave.

Dias, Pedro. 2002. Diu em 1634. Documentos e notas para um retrato de uma praça portuguesa no Guzarate. Academia de Marinha, Lisbon.

Dik, Simon C. 1997. *The theory of Functional Grammar*, volume 1. Berlin: Mouton de Gruyter, 2nd edition.

Dixon, Robert Malcolm Ward. 1979. Ergativity. *Language* 55:59–138.

Dixon, Robert Malcolm Ward. 1982. *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton de Gruyter.

Dixon, Robert Malcolm Ward. 2008. Comparative constructions. *Studies in Language* 32:787–817.

Emeneau, Murray B. 1956. India as a linguistic area. *Language* 32:3–16.

Estado da Índia - Direcção dos Serviços de Economia. 1961. *Anuário Estatístico 1957*. Goa: Imprensa Nacional.

Evans, Nicholas. 2000. Word classes in the world's languages. In *Morphologie, Ein internationales Handbuch zur Flexion und Wortbildung*, ed. Geert Booij, Christian Lehmann, and Joachim Mugdan, 708–732. Berlin: Mouton de Gruyter.

de Federici, Cesare. 1588. *The voyage and trauaile: of M. Cæsar Frederick, merchant of Venice, into the East India, the Indies, and beyond the Indies*. London: printed by Richard Jones and Edward White.

Ferraz, Luís Ivens. 1987. Portuguese creoles of West Africa and Asia. In *Pidgin and Creole*

languages, ed. Glenn Gilbert. Honolulu: University of Hawaii Press.

Fitch, Ralph. 1599. *The voyage of Mr. Ralph Fitch merchant of London begunne in 1583, and ended 1591*. London.

Foley, William A. 2005. Personhood and linguistic identity, purism and variation. In *Language description and documentation*, ed. Peter K. Austin, volume 3, 157–180. London: SOAS.

Furtad, Dacho. 1999. *Concanim-Inglez dicionar: Konkani-English Dictionary*. New Delhi/Madras: Asian Educational Services.

Gil, David. 2004. Riau Indonesian *Sama*: Explorations in macrofunctionality. In *Coordinating constructions*, ed. Martin Haspelmath, 371–424. Amsterdam: John Benjamins.

Giles, Howard. 1977. *Language, ethnicity, and intergroup relations*. New York: Academic Press.

Giles, Howard, R. Bourhis, and D. Taylor. 1977. Towards a theory of language in ethnic group relations. In *Language, ethnicity, and intergroup relations*, ed. Howard Giles. New York: Academic Press.

Gumperz, John J., and Robert Wilson. 1971. Convergence and creolization: A case from the Indo-Aryan/Dravidian border. In *Pidginization and creolization of languages*, ed. Dell H. Hymes, 251–273. London: Cambridge University Press.

Hancock, Ian F. 1975. Malacca Creole Portuguese: Asian, African or European? *Anthropological Linguistics* 17:211–236.

Haspelmath, Martin. 1997. *Indefinite pronouns*. Oxford: Oxford University Press.

Haspelmath, Martin. 2003. The geometry of grammatical meaning: semantic maps and crosslinguistic comparison. In *The new psychology of language*, ed. Michael Tomasello, volume 2, 211–43. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Haspelmath, Martin. 2004. Coordinating constructions; an overview. In *Coordinating constructions*, ed. Martin Haspelmath, 3–39. Amsterdam: John Benjamins.

Haspelmath, Martin. 2005a. Indefinite pronouns. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 46. Oxford: Oxford University Press. Available online at http://wals.info/feature/46.

Haspelmath, Martin. 2005b. 'want' complement clauses. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 124. Oxford: Oxford University Press. Available online at http://wals.info/feature/124.

Haspelmath, Martin. 2007a. Ditransitive alignment splits and inverse alignment. *Functions of Language* 14:79–102.

Haspelmath, Martin. 2007b. Pre-established categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11:119–132.

Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26:210–231.

Heine, Bernd, and Tania Kuteva. 2002. *World lexicon of grammaticalization*. Cambridge: Cambridge University Press.

Hengeveld, Kees. 1992a. *Non-verbal predication; theory, typology, diachrony*. Berlin: Mouton de Gruyter.

Hengeveld, Kees. 1992b. Parts of Speech. In *Layered structure and reference in a func-*

tional perspective, ed. Michael Fortescue, Peter Harder, and Lars Kristofferson, 29–56. Amsterdam: John Benjamins.

Hengeveld, Kees, and J. Lachlan Mackenzie. 2008. *Functional discourse grammar*. Oxford: Oxford University Press.

Hengeveld, Kees, Jan Rijkhoff, and Anna Siewierska. 2004. Parts-of-speech systems and word order. *Journal of Linguistics* 40:527–570.

van Hoecke, W. 1996. The Latin dative. In *The dative*, ed. William van Belle and Willy van Langendonck, volume 1, 79–124. Amsterdam: John Benjamins.

Jackson, Kenneth David. 1990. *Sing without shame. Oral traditions in Indo-Portuguese Creole verse*. Amsterdam: Benjamins.

Jayaseelan, K. A. 1990. The Dative Subject Construction and the Pro-drop parameter. In *The notion of subject in south asian languages*, ed. Mahendra K. Verma and K. P. Mohanan, 269–284. Madison: University of Wisconsin.

Jayaseelan, K. A. 2004. The possessor-experiencer dative in Malayalam. In *Non-nominative subjects*, ed. Peri Bhaskararao and Karumuri Venkata Subbarao, volume 1, 227–244. Amsterdam: John Benjamins.

Jayasuriya, Shihan de Silva. 1996. Indo-Portuguese Songs of Sri Lanka: The Nevill Manuscript. *Bulletin of the School of Oriental and African Studies* 59:253–267.

Jayasuriya, Shihan de Silva, and Richard Pankhurst. 2003. On the African diaspora in the Indian Ocean. In *The african diaspora in the indian ocean*. New Jersey: Africa World Press.

Joos, Martin, ed. 1957. *Readings in linguistics. The development of descriptive linguistics in America since 1925*. American Council of Learned Societies.

Jourdan, Christine. 1991. Pidgins and creoles: the blurring of categories. *Annual Review of Anthropology* 20:187–209.

Kachru, Braj B., Yamuna Kachru, and S. N. Sridhar, ed. 2008. *Language in South Asia*. Cambridge: Cambridge University Press.

Kachru, Yamuna. 1982. Conjunct verbs in Hindi-Urdu and Persian. In *South asian review: studies in south asian languages & linguistics*, ed. P. J. Mistry, 117–126. Jacksonvilee, FL: South Asian Literary Association.

Kachru, Yamuna. 2008. Hindi-Urdu-Hindustani. In *Language in south asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 81–102. Cambridge: Cambridge University Press.

Karcevski, Serge. 1969a. Introduction à l'étude de l'interjection. In *A Geneva School Reader*, ed. Robert Godel. Bloomington, IN: Indiana University Press.

Karcevski, Serge. 1969b. Sur la parataxe et la syntaxe en russe. In *A geneva school reader*, ed. Robert Godel. Bloomington, IN: Indiana University Press.

Khanina, Olesya. 2008. How universal is *wanting*? *Studies in Language* 32:818–865.

Koul, Omkar N. 2003. Kashmiri. In *The Indo-Aryan languages*, ed. George Cardona and Dhanesh Jain, 895–952. New York: Routledge.

Labov, William. 1966. *The Social Stratification of English in New York City*. Washington, D.C.: Center for Applied Linguistics.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania

Press.

Lambert, H. M. 1971. *Gujarati language course*. Cambridge: Cambridge University Press.

Laval, Pyrard de. 1998. *Voyage de Pyrard de Laval aux Indes Orientales (1601-1611)*. Paris: Chandeigne.

Lefebvre, Claire. 1998. *Creole genesis and the acquisition of grammar. the case of Haitian*. Cambridge: Cambridge University Press.

Lefebvre, Claire. 2004. *Issues in the study of Pidgin and Creole languages*. Amsterdam: Benjamins.

Lehiste, Ilse. 1988. *Lectures on language contact*. Cambridge, Mass.: MIT Press.

Leão, Duarte Nunes de. 1576. *Orthographia da lingoa portuguesa*.

Leão, Mário César. 1996. *A Província do Norte do Estado da Índia*. Macau: Instituto Cultural de Macau.

LePage, Robert Brock, and Andrée Tabouret-Keller. 1985. *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Lodhi, Abdulaziz Y. 1992. African settlements in India. *Nordic Journal of African Studies* 1:83–86.

Lopes, D. 1936. *A expansão da língua portuguesa no Oriente, durante os séculos XVI, XVII e XVIII*. Barcelos: Portucalense Editora.

Lourenço, Inês. 2005. Diu: espaço de reencontro. *Revista Oriente* 56–70.

Luís, Ana R. 2008. Tense marking and inflectional morphology in Indo-Portuguese creoles. In *Roots of Creole Structures: Weighing the contribution of substrates and superstrates*, ed. Susanne Michaelis, 83–121. Amsterdam: John Benjamins.

Mackenzie, J. Lachlan. in press. Content interrogatives in a sample of 50 languages. *Lingua* .

Mahajan, Anoop. 2004. On the origin of non-nominative subjects. In *Non-nominative subjects*, ed. Peri Bhaskararao and Karumuri Venkata Subbarao, volume 1, 293–299. Amsterdam: Benjamins.

Malchukov, Andrej L. 2004. Towards a semantic typology of adversative and contrast marking. *Journal of Semantics* 21:177–198.

Malchukov, Andrej L., Martin Haspelmath, and Bernard Comrie. forthcoming. Ditransitive constructions: a typological overview. In *Studies in ditransitive constructions*, ed. Andrej Malchukov, Martin Haspelmath, and Bernard Comrie. Berlin: Mouton de Gruyter.

Masica, Colin P. 1976. *Defining a linguistic area: South Asia*. Chicago: University of Chicago Press.

Masica, Colin P. 1991. *The Indo-Aryan languages*. Cambridge: Cambridge University Press.

Mateus, Maria Helena Mira, Ana Maria Brito, Inês Duarte, Isabel Hub Faria, Sónia Frota, Gabriela Matos, Fátima Oliveira, Marina Vigário, and Alina Villalva. 2003. *Gramática da Língua Portuguesa*. Lisbon: Caminho, 6th edition.

Mateus, Maria Helena Mira, and Ernesto d'Andrade. 2000. *The phonology of Portuguese*. Oxford: Oxford University Press.

Matos, Artur Teodoro de. 1999. Prefácio. In *O Tombo de Diu, 1592*. Lisbon: Comissão Nacional para as Comemorações dos Descobrimentos Portugueses.

Matos, L. 1968. O português - língua franca no Oriente. In *Colóquios sobre as províncias do Oriente*, ed. A. Moreira. Lisbon: Junta de Investigações Ultramarinas.

Matos, Manuel Cadafaz de. 1987. Para uma análise sistémica das inter-relações do Português e do «Gujaráti» no território de Damão (através de espécimes folclóricos ali recolhidos). In *Congresso sobre a situação actual da Língua Portuguesa no mundo*, volume 2, 323–349. Lisbon: Instituto de Cultura e Língua Portuguesa.

McMahon, April M. S. 1994. *Understanding language change*. Cambridge: Cambridge University Press.

McWhorter, John. 1998. Identifying the creole prototype: vindicating a typological class. *Language* 74:788–818.

Miestamo, Matti. 2005. Symmetric and Asymmetric Standard Negation. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 113. Oxford: Oxford University Press. Available online at http://wals.info/feature/113.

Milroy, Lesley. 1987. *Observing and analysing natural language: a critical account of sociolinguistic method*. Oxford: Basil Blackwell.

Mistry, P. J. 1997. Gujarati phonology. In *Phonologies of Asia and Africa (including the Caucasus)*, ed. Alan S. Kaye and Peter T. Daniels, volume 2, 653–673. Winona Lake, IN: Eisenbrauns.

Mistry, P. J. 2004. Subjecthood of non-nominatives in Gujarati. In *Non-nominative subjects*, ed. Peri Bhaskararao and Karumuri Venkata Subbarao, volume 2, 1–32. Amsterdam: John Benjamins.

Mithun, Marianne. 1995. On the relativity of Irreality. In *Modality in grammar and discourse*, ed. Joan Bybee and Suzanne Fleischman, 367–388. Amsterdam: Benjamins.

Modi, Bharati. 1994. The phonetics and phonology of mid vowels in Gujarati. *Indian Linguistics* 55:51–76.

Mohanan, Tara. 2006. Grammatical verbs (with special reference to light verbs). In *The Blackwell companion to syntax*, ed. Martin Everaert and Henk van Riemsdijk, volume 2, chapter 31, 459–492. Malden, MA: Blackwell.

Morais, Carlos Alexandre de. 1997. *Cronologia geral da Índia Portuguesa. 1498-1962*. Lisbon: Editorial Estampa.

Moravcsik, Edith. 2003. A semantic analysis of associative plurals. *Studies in Language* 27:469–503.

Mosel, Ulrike. 2004. Dictionary-making in endangered language communities. In *Language documentation and description*, ed. Peter K. Austin, volume 2, 39–54. London: SOAS.

Moura, João Herculano Rodrigues de. 1901. *Relatório sobre a administração e serviços do governo do districto de Diu, anno de 1900*. Nova Goa: Imprensa Nacional.

Mufwene, Salikoko S. 1996. The founder principle in creole genesis. *Diachronica* 13:83–134.

Mufwene, Salikoko S. 2001. *The ecology of language evolution*. Cambridge: Cambridge University Press.

Myers-Scotton, Carol. 1993. *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press.

Newitt, Malyn. 2005. *A history of Portuguese overseas expansion 1400-1668*. London:

Routledge.

Nunes, Antonio Raphael Pereira. 1898. *Relatório sobre os serviços do Governo do districto de Diu: relativo aos anos de 1896 e 1897*. Nova Goa: Imprensa Nacional.

Oliveira, Fernão de. 1536. *Grammatica da lingoagem portuguesa*.

Orta, Garcia de. 1563. *Colóquios dos simples e drogas he cousas medicinais da Índia e assi dalgũas frutas achadas nella onde se tratam algũas cousas tocantes a medicina, pratica, e outras cousas boas pera saber*. Goa.

Pais, Francisco. 1592. Tombo de Diu. Manuscript kept at *Goa Historical Archives*, doc. 624 - 'Tombo de Diu'.

Palmié, Stephan. 2006. Creolization and its discontents. *Annual Review of Anthropology* 35:433–456.

Pandharipande, Rajeshwari. 2003. Marathi. In *The Indo-Aryan languages*, ed. George Cardona and Dhanesh Jain, 698–728. New York: Routledge.

Pandit, Prabodh B. 1955. *E* and *o* in Gujarati. *Indian Linguistics* 14:36–44.

Pandit, Prabodh B. 1979. Perspectives on sociolinguistics in India. In *Language and society: Anthropological issues*, ed. W. C. McCormick and S. A. Wurm. The Hague: Mouton.

Pankhurst, Richard. 2003. The Ethiopian diaspora to India: the role of Habshis and Sidis from medieval times to the end of the eighteenth century. In *The African Diaspora in the Indian Ocean*, ed. Shihan de Silva Jayasuriya and Richard Pankhurst. New Jersey: Africa World Press.

Parkvall, Mikael. 2002. Cutting off the branch. In *Pidgin and Creole linguistics in the twenty-first century*, ed. Glenn Gilbert, 355–368. New York: Peter Lang.

Payne, Thomas E. 1997. *Describing morphosyntax*. Cambridge: Cambridge University Press.

Pearson, M. N. 1976. *Merchants and rulers in Gujarat; the response to the Portuguese in the sixteenth century*. Berkeley/Los Angeles: University of California Press.

Pinto, Jeanette. 1992. *Slavery in Portuguese India (1510-1842)*. Bombay: Himalaya Publishing House.

Pissarra, José Virgílio Amaro. 2002. *Chaul e Diu; 1508 e 1509*. Lisbon: Tribuna.

Quadros, Jeronymo. 1899. *Diu, apontamentos para sua historia e chorographia*. Nova Goa: Tipographia Fontainhas.

Quadros, Jeronymo. 1907. *Cartas de Diu*. first series (1902-1905). Nova Goa: Tipographia Fontainhas.

Ramat, Paolo, and Davide Ricca. 1994. Prototypical adverbs: on the scalarity/radiality of the notion of ADVERB. *Rivista di Linguistica* 6:289–326.

Rijkhoff, Jan. 2000. When can a language have adjectives? In *Approaches to the typology of word classes*, ed. Petra M. Vogel and Bernard Comrie, 217–257. Berlin: Mouton de Gruyter.

Rijkhoff, Jan. 2002. *The noun phrase*. Oxford: Oxford University Press.

Rijkhoff, Jan. 2008. On flexible and rigid nouns. *Studies in Language* 32:727–752.

Romaine, Suzanne. 1995a. *Bilingualism*. Oxford: Blackwell, 2nd edition.

Romaine, Suzanne. 1995b. The grammaticalization of Irrealis in Tok Pisin. In *Modality in*

grammar and discourse, ed. Joan Bybee and Suzanne Fleischman, 389–427. Amsterdam: Benjamins.

Ross, Malcolm D. 1996. Contact-induced change and the comparative method: Cases from Papua New Guinea. In *The comparative method reviewed: Regularity and irregularity in language change*, ed. Mark Durie and Malcolm D. Ross, 180–217. New York: Oxford University Press.

Ross, Malcolm D. 2001. Contact-induced change in Oceanic languages in North-West Melanesia. In *Areal diffusion and genetic inheritance*, ed. Alexandra Y. Aikhenvald and R. M. W. Dixon, 134–166. Oxford: Oxford University Press.

Rougé, Jean-Louis. 2004. *Dictionnaire étymologique des créoles portugais d'Afrique*. Karthala.

Sá, Francisco Xavier Valeriano de. 2004. *O padroado português do Oriente e os mitrados da Sé de Goa*. Lisbon: Plátano Editora.

Sasse, Hans-Jürgen. 1993. Syntactic categories and subcategories. In *Syntax: Ein internationales Handbuch zeitgenössischer Forschung*, ed. Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, 646–686. Berlin: Walter de Gruyter.

Saxena, Anju. 2005. It takes two to tango: linguistic and cultural co-variation in digital documentation. In *Language documentation and description*, ed. Peter K. Austin, volume 3, 181–195. London: SOAS.

Schachter, Paul. 1985. Parts-of-Speech systems. In *Language typology and syntactic description*, ed. Timothy Shopen, 3–61. Cambridge: Cambridge University Press.

Schuchardt, Hugo. 1883. Kreolische Studien III. Über das Indoportugiesische von Diu. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (philosophisch-historische Klasse)* 103:3–18.

Schuchardt, Hugo. 1890. Kreolische Studien IX. Ueber das Malaioportugiesische von Batavia und Tugu. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften zu Wien (Philosophisch-historische Klasse)* 122:1–258.

Schuchardt, Hugo. 1899. Beiträge zur Kenntnis des creolischen Romanisich. V. Allgemeineres über das Indoportugiesische (Asioportugiesische). *Zeitschrift für Romanische Philologie* 13:476–516.

Scovel, Thomas. 1998. *Psycholinguistics*. Oxford: Oxford University Press.

Selkirk, Elisabeth. 1984. On the major class features and syllable theory. In *Language sound structure*, ed. Mark Aronoff and Richard T. Oehrle, 107–136. Cambridge, MA: MIT Press.

Shackle, Christopher. 2003. Panjabi. In *The Indo-Aryan languages*, ed. George Cardona and Dhanesh Jain, 581–621. New York: Routledge.

Shokoohy, Mehrdad. 2003. The Zoroastrian fire temple in the ex-Portuguese colony of Diu, India. *Journal of the Royal Asiatic Society* 13:1–20.

Silva, Rosa Virgínia Mattos e. 1989. *Estruturas trecentistas. elementos para uma gramática do Português Arcaico*. Lisbon: Imprensa Nacional - Casa da Moeda.

Silva, Rosa Virgínia Mattos e. 1991. *O português arcaico: fonologia*. São Paulo: Contexto.

Singh, K. S., B. R. Solanki, N. K. Sinha, and Jaime F. Pereira, ed. 1994. *People of India: Daman and Diu*, volume XIX. Bombay: Popular Prakashan / Anthropological Survey

of India.

Smith, Ian. 1977. Sri Lanka Creole Portuguese Phonology. Doctoral Dissertation, Cornell University, Ithaca NY.

Smith, Ian. 1984. The development of morphosyntax in Sri Lanka Portuguese. In *York papers in Linguistics*, ed. Mark Sebba and Loreto Todd. York: University of York.

Smith, Ian. 2008a. Pidgins, Creoles and Bazaar Hindi. In *Language in South Asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 253–268. Cambridge: Cambridge University Press.

Smith, Norval. 1995. An annotated list of creoles, pidgins, and mixed languages. In *Pidgins and Creoles - an introduction*, ed. Jacques Arends, Pieter Muysken, and Norval Smith, 331–374. Amsterdam: John Benjamins.

Smith, Norval. 2008b. Creole phonology. In *The Handbook of Pidgin and Creole Studies*, ed. Silvia Kouwenberg and John Victor Singler, 98–129. Malden: Wiley-Blackwell.

Southworth, Franklin C. 2005. *Linguistic archeology of South Asia*. London: Routledge-Curzon.

Sridhar, S. N. 2008. Language contact and convergence in South Asia. In *Language in South Asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 234–252. Cambridge: Cambridge University Press.

Stassen, Leon. 1985. *Comparison and Universal Grammar*. Oxford: Basil Blackwell.

Stassen, Leon. 2000. AND-languages and WITH-languages. *Linguistic Typology* 4:1–54.

Stassen, Leon. 2005a. Nominal and verbal conjunction. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 64. Oxford: Oxford University Press. Available online at http://wals.info/feature/64.

Stassen, Leon. 2005b. Noun phrase coordination. In *The World Atlas of Language Structures*, ed. Martin Haspelmath, Matthew S. Dryer, David Gil, and Bernard Comrie, chapter 63. Oxford: Oxford University Press. Available online at http://wals.info/feature/63.

Subbarao, Karumuri Venkata. 2008. Typological characteristics of South Asian languages. In *Language in South Asia*, ed. Braj B. Kachru, Yamuna Kachru, and S. N. Sridhar, 49–78. Cambridge: Cambridge University Press.

Teyssier, Paul. 2001. *História da Língua Portuguesa*. Lisboa: Livraria Sá da Costa Editora, 8th edition.

Teza, E. 1872. Indoportoghese. *Studi Filologici* 5.

The International Phonetic Association. 1999. *Handbook of the International Phonetic Association. A guide to the use of the International Phonetic Alphabet*. Cambridge: Cambridge University Press.

Theban, L. 1985. *Situação e perspectivas do português e dos crioulos de origem portuguesa na Índia e no Sri Lanka*. Lisbon: Instituto de Cultura e Língua Portuguesa.

Thomason, Sarah Grey, and Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley and Los Angeles: University of California Press.

Tibbetts, G. R. 1971. *Arab navigation in the indian ocean before the coming of the portuguese*. London: The Royal Asiatic Society of Great Britain and Ireland.

Tisdall, William St. Clair. 1892. *A simplified gramar of the Gujarati language, together*

with a short reading book and vocabulary. London: K. Paul, Trench, Trubner.

Tomás, Maria Isabel. 1992. *Os crioulos portugueses do oriente; uma bibliografia*. Macau: Instituto Cultural de Macau.

Trudgill, Peter, and Jean Hannah. 2002. *International English: A guide to the varieties of Standard English*. Oxford: Oxford University Press, 4th edition.

Turner, Ralph L. 1921. Gujarati Phonology. *Journal of the Royal Asiatic Society* 505–544.

Turner, Ralph L. 1925. The *e* and *o* vowels in Gujarātī. *Ashutosh Mukherji Silver Jubilee Volume, Orientalia* II:337–347. [reprinted in Turner 1973, pp. 229-238].

Turner, Ralph L. 1973. *Indo-Aryan linguistics - Collected papers 1912-1973*. School of Oriental and African Studies.

Vardarjan, Lotika. 1989. Diu and its trade. *Studia* 49:365–378.

Vasconcellos, José Leite de. 1892. *Sur le dialecte portugais de Macao*. Lisboa: Imprensa Nacional.

Vasconcellos, José Leite de. 1901. Esquisse d'une dialectologie portugaise. Doctoral Dissertation, Université de Paris.

Wierzbicka, Anna. 2000. Lexical prototypes as a universal basis for cross-linguistic identification of 'parts of speech'. In *Approaches to the typology of word classes*, ed. Petra M. Vogel and Bernard Comrie, 285–317. Berlin: Mouton de Gruyter.

Xavier, Carlos. 1981. Um ataque francês a Diu. *Revista da Academia da Língua e Cultura Portuguesa* 1:3–11.

Özbaran, Salih. 2001. Ottomans as 'Rumes' in Portuguese Sources in the Sixteenth Century. *Portuguese Studies* 17:64–74.

Appendix - texts

STORY 1: This story was recorded on March 3rd, 2005, and it was told by a male student aged about 8 outside the *Nirmala Mata* Catholic School in Diu Town.

ũ   di   ũ   rey   də jungle *t-iŋ,   i   ũ   piken rat   t-iŋ   vay   nə   d-el*
one day one king of jungle EXS-PST and one small mouse IPFV-PST go.INF LOC of-3s

kamiŋ. i   rey   də jungle *fal-o   use ɛ   tãt piken, use uki   a*
way   and king of jungle say-PST 2s   COP$_i$.NPST so small   2s   what IRR.NPST

faz-e? i   ũ   jungly jat vey   i   amər-o pɛ d-ikəl   lion. ɛ-r
do-INF and one jungly jat come.PST and tie-PST leg of-DEM$_d$ lion   COP$_i$-PST

pə   kum-e   a   el. i   rat   kir-iŋ   salv-a   a   el. i   rat   foy,
PURP eat-INF DAT 3s and mouse want-PST save-INF DAT 3s and mouse go.PST

murde-w murde-w pikəl-iŋ, sult-o.   i   dəpəy lion fal-o   ki   muyt ubrigad.
bite-PST bite-PST bit-DIM release-PST and after lion say-PST CMP much thanks

istɔr kab-o.
story finish-PST

'Once upon a time there was a king of the jungle, and a little mouse was going about his way. And the king of the jungle told him: "you are so small, what are you going to do?". And then a wild man came and bound the lion's leg, he intended to eat him. And the mouse wanted to save him. So the mouse went and gnawed and gnawed [at the rope] and set him free. And then the lion said: "thank you very much". The story is finished.'

STORY 2: This story was recorded on March 3rd, 2005, and it was told by a male student aged about 9 outside the *Nirmala Mata* Catholic School in Diu Town.

ũ   makak   t-iŋ   i   ũ   crocodile. *ikəl   makak   t-iŋ   vay,   el*
one monkey EXS-PST and one crocodile DEM*d* monkey IPFV-PST go.INF 3s

t-iŋ   vay   nə   ũ jungle *pu   traz-e   koys, aroz tud, ali   t-iŋ*
IPFV-PST go.INF LOC one jungle PURP bring-INF thing rice SIML there EXS-PST

ũ   loj, el foy   traz-e   aroz i   ali   t-iŋ   mar, el vi-w   ũ
one shop 3s go.PST bring-INF rice and there EXS-PST sea   3s see-PST one

crocodile *nə?   sənãw* crocodile *dis   use ɔn   vay?   el dis   ki   yo*
crocodile REQ then   crocodile say.PST 3s   where go.NPST 3s say.PST CMP 1s

vay   agɔr nu   kõpr-a   aroz, dəpəy də aroz yo vay   nə   kazəmẽt də
go.NPST now LOC buy-INF rice   after of rice 1s go.NPST LOC wedding of

mĩ   irmã. dəpəy ikəl   crocodile dis   use vay,   dəpəy də trey di   use
1s.POSS sister after DEM*d* crocodile say.PST 2s   go.NPST after of three day 2s

vẽy   volta-d.   kɔm   bẽy i   vẽy.   el foy,   el fik-o
come.NPST return-PTCP eat.NPST well and come.NPST 3s go.PST 3s become-PST

med, el kõpr-o   aroz i   foy   ali   kaz   də irmã. dez di el kume-w, dəpəy
fear 3s buy-PST rice and go.PST there house of sister ten day 3s eat-PST after

el foy   i   el dis   pə irmã irmã asĩ ɛ   ki   ũ   crocodile yo
3s go.PST and 3s say.PST DAT sister sister thus COP*i*.NPST CMP one crocodile 1s

apiŋ-o   i   el kɛr   kum-e a   mĩ   nə   istrad. d-el irmã dis
meet-PST and 3s want.NPST eat-INF DAT 1s.OBL LOC road   of-3s sister say.PST

326

nã   *faz*   *mal,*   *tɔm*   *es*   *daba,* *use* *ĕtr-a*   *dẽt*   *d-es*   *daba,*
NEG$_{cl}$ make.NPST wrong take.NPST DEM$_p$ tin   2s   enter-INF inside of-DEM$_p$ tin

yo tə   *faz-e*   *burak,* *use a*   *ãd-a.*   *dəpəy el foy,*   *el foy*
1s IPFV.NPST make-INF hole   2s   IRR.NPST walk-INF after   3s go.PST 3s go.PST

i   crocodile *vey.*   *el dis.*   *nãw,* *ikəl*   *el nã*   *uz-o.*   *el vey,*
and crocodile come.PST 3s say.PST NEG$_p$ DEM$_d$ 3s NEG$_{cl}$ use-PST 3s come.PST

makak *vey*   *i*   crocodile *dis*   *yo kɛr*   *kum-e* *d-use kurəsãw, d-use*
monkey come.PST and crocodile say.PST 1s want.NPST eat-INF of-2s   heart   of-2s

kurəsãw ɔn   *te?*   *dəpəy makak* *dis*   *así̃* *mĩ̃*   *kurəsãw te*
heart   where EXS.NPST after moneky say.PST thus 1s.POSS heart   EXS.NPST

mĩ̃   *kaz*   *ali,*   *kaz*   *də mĩ̃*   *irmã, yo larg-o*   *kurəsãw i*   *vey,*
1s.POSS house there house of 1s.POSS sister 1s drop-PST heart   and come.PST

use vay   *i*   *trag.*   *dəpəy* crocodile *foy*   *ali*   *kaz*   *də irmã, foy*
2s go.NPST and bring.NPST after crocodile go.PST there house of sister go.PST

ver   *pu*   *ikəl*   *kurəsãw. ikəl*   *makak*   *dis*   *use vay*   *i*   *ver,*
see.INF DAT DEM$_d$ heart   DEM$_d$ monkey say.PST 2s   go.NPST and see.INF

use vay   *i*   *ver*   *kaz*   *də mĩ̃*   *irmã, makak*   *dis*   *pu*   crocodile.
2s go.NPST and see.INF house of 1s.POSS sister monkey say.PST DAT crocodile

crocodile *foy*   *i*   *ver,*   *el dis*   *ki*   *kurəsãw d-ikəl*   *makak*   *ɔn*
crocodile go.PST and see.INF 3s say.PST CMP heart   of-DEM$_d$ monkey where

te?   *ikəl*   *irmã dis*   *kurəsãw də makak* *kwɔlki di*   *ad*
EXS.NPST DEM$_d$ sister say.PST heart   of monkey any   day IRR.NPST

sa-i   *fɔr?* *kurəsãw də makak*   *dẽt*   *d-el* *mem*   *korp.* *dəpəy el fik-o*
leave-INF out heart   of monkey inside of-3s EMPH body after 3s become-PST

ĩ̃truja-d   *i*   *el foy*   *kaz.*   *i*   *akab-o*   *istɔr.*
deceive-PST and 3s go.PST house and finish-PST story

‘There was monkey and a crocodile. The monkey was going into a jungle to bring some
things, rice and all that, there was a shop there so he went to buy rice; and there was also
the sea, he saw a crocodile, you see? Then the crocodile said: "Where are you going?". He
said: "I am now going to buy rice, and after that I go to my sister's wedding". Then the
crocodile said: "Go and come back in three days. Eat well and return". So he went, he
was scared, he bought the rice and he went to his sister's house. He ate for ten days, and

then he went and he said to his sister: "Sister, this has happened; I ran into a crocodile and he wants to eat me on my way home". His sister said: "Don't worry. Take this tin, you will enter this tin, I will make a hole in it and you can walk". Then he went, he went and the crocodile came. No, he didn't use that. He came, the monkey came and the crocodile said: "I want to eat your heart, where's your heart?". Then the monkey said this: "My heart is at home over there, in my sister's house, I dropped the heart and came here, you go and bring it". Then the crocodile went to the sister's house, he went to check on the heart. The monkey said: "Go and see"; "go and see in my sister's house", the monkey told the crocodile. The crocodile went to check, he said: "Where's the monkey's heart?". And the sister said: "Will the monkey's heart ever come out? The monkey's heart is inside his own body". Then he felt deceived and he went home. And the story is finished.'

STORY 3: This very short story was recorded on March 3rd, 2005, and it was told by a male student aged about 7 outside the *Nirmala Mata* Catholic School in Diu Town.

ũ   di   t-iŋ   ũ   crow. el kume-w, el kume-w i   dəpəy el kɛri-w   ag.
one day EXS-PST one crow 3s eat-PST 3s eat-PST and after 3s want-PST water

dəpəy atər-o   pɛd   nə   ag,   dəpəy atər-o   dez pɛd   kavok, atər-o
after throw-PST stone LOC water after throw-PST ten stone cave   throw-PST

pɛd   dəpəy bebe-w   ag.
stone after   drink-PST water

'There was a crow once. He ate, he ate and then he wanted water. Then he threw stones in the water, then he threw ten stones into the hole, he threw the stones and then he drank water.'

DIALOGUE: This dialogue excerpt was recorded on January 28th, 2007. The participants are the researcher (whose interventions are given in square brackets and left unanalysed) and a middle-aged lady whose speech revealed certain SP traits (contrast, for instance, this dialogue with the discussion in section 7.1.1). The conversation took place at the lady's house in Diu Town.

kwɔn vey   dumĩg pasa-d,   dumĩg pasa-d   kwɔn ista-v   aki
when come.PST sunday pass-PTCP sunday pass-PTCP when EXS-IPFV here

ɛ-r   mĩ   an. yo nã   di-s   pə   nĩge, yo fez   así surprise
COP$_s$-PST 1s.POSS year 1s NEG$_{cl}$ say-PST DAT nobody 1s make.PST thus surprise

an   así. də maner, yo t-iŋ   doy võtad: ker-iŋ   faz-e   an, fɛstej-a
year thus of manner 1s have-PST two wish   want-PST make-INF year celebrate-INF

an, nã   ker-ɔŋ   fɛstej-a   an. ultim mem   d-ɔr,   dəpəy də afternoon
year NEG$_{cl}$ want-PST celebrate-INF year last   EMPH of-hour after of afternoon

pẽs-o   bam   faz-e   ũ   piken koyz dẽt   də famil. means yo nã
think-PST HORT make-INF one small thing inside of family means 1s NEG$_{cl}$

te   muyt famil, doyz irmã, subriŋ, mĩ   fil   i   nɛt.   v-iw
have.NPST much family two sister nephew 1s.POSS child and grandchild see-PST

n-ikəl   di kẽ t-iŋ   mĩ   kaz, sɔ   ikəl tãt   piso.
LOC-DEM$_d$ day who EXS-PST 1s.POSS house only DEM$_d$ so.much person

dəpəy di-s   ki   faz   nə   d-use ɔm vay   fik-a   mufin d-es.
after   say-PST CMP make.NPST REQ of-2s man go.NPST become-INF sad   of-DEM$_d$

yo foy   yo fɛstej-o   así d-əkəl   maner. time atɛ   dɛz i   me akəb-o
1s go.PST 1s celebrate-PST thus of-DEM$_d$ manner also until ten and half finish-PST

330

fɛs.
party

[dɛz i me dəpəy akəbo?]

atɛ   dɛz, dɛz i   me akəb-o.   jə   ɛ   tẽp də fri   təme nə?
until ten   ten and half finish-PST already COP$_i$.NPST time of cold also   REQ

kriãs təme fik,   ɛl   fik   muyt lõj ali   nə   OIDC.
child also   stay.NPST 3fs stay.NPST much far there LOC OIDC

[kẽ fik ali nə OIDC?]

nɛt.   fil   də mĩ   ku   nɛt,   ɛl   fik   ali   nə   OIDC.
grandchild child of 1s.OBL COM grandchild 3fs stay.NPST there LOC OIDC

də maner   dəpəy pə   vay   suziŋ təme, yo di-s   use akab   kum-e
of manner after   DAT go.INF alone also   1s say-PST 2s finish.NPST eat-INF

lɔg   i   vay   ĩmɔr use. i   pas-o   time asĩ, nɔs kwɔn fik-o
immediately and go.NPST away 2s   and pass-PST time thus 1p   when become-PST

õz   ɔr   nã   fik-o   sabe-n.   dəpəy durmi-w, maŋa   lɛŭt-o,
eleven hour NEG$_{cl}$ become-PST know-PROG after   sleep-PST morning rise-PST

foy   trabay.
go.PST work

'When you came last sunday, last sunday when you were here it was my birthday. I didn't tell anyone, I made like a surprise brithday. so, I was of two minds: I wanted to make a birthday (party), to celebrate my birthday, I didn't want to celebrate my birthday. At the very last moment, after noon, I though "let's prepare a small thing inside the family". I mean, I don't have much family, two sisters, nephews, my children and grandchildren. That day you saw who was in my house, right? Only those people. And then they said "Come on, do it, otherwise your husband will be sad". And went and I celebrated in that manner. But at half past ten it was over.

[Half past ten, then it was over?]

At ten, ten thirty is was over. It's already the cold season, you see? And daughter lives, she lives very far, at the OIDC.

[Who lives at the OIDC?]

My grandchildren. My daughter and the grandchildren, she lives at the OIDC. So also to go home alone, I said: "Finish eating quickly and go away". So we passed the time, we didn't even notice went it was eleven o'clock. Then (I) slept, in the morning (I) woke up and went to work.'

Samenvatting (Dutch summary)

Het kleine eiland Diu, gesitueerd op het zuidelijkste eindje van het schiereiland Saurashtra te Gujarat (India) met een strategisch uitzicht over de Arabische Zee, wekte in haar geschiedenis een schijnbaar buitensporige koloniale interesse. Een van de vele gevolgen van haar lange overheersing door het ver weg gelegen Portugal (1535-1961) was de vorming van een locale variant van het Indo-Portugees, een contacttaal die ontstond door de intersectie van diverse talige invloeden, met name het Gujerati en het Portugees. Alhoewel de uit het Portugees stammende Aziatische creolen al vanaf het einde van de negentiende eeuw wetenschappelijk bestudeerd worden, is de nauwkeurige taalkundige beschrijving van deze talen een betrekkelijk recente trend. Dit onderzoek biedt een taalkundige documentatie van het moderne Indo-Portugees van Diu en situeert deze taal daarbij uitdrukkelijk binnen een reconstructie van de historische en sociaaldemografische context. De bedoeling van dit onderzoek is bij te dragen aan het ontluikende begrip van de vorming, ontwikkeling en actuele vitaliteit van contacttalen in (Zuid)Azië en elders.

Dit proefschrift bestaat uit drie aparte onderdelen. Deel I, 'Achtergrond' ['Background'], geeft alle niet-taalkundige informatie die nodig is om de vorming en ontwikkeling van het Indo-Portugees op Diu te begrijpen. Hoofdstuk 1 is daarbij een korte introductie. Het volgende hoofdstuk biedt een synchrone analyse van de omstandigheden waarin het Indo-Portugees op het eiland ontstond. Het gaat daarbij zowel in op de bevolkingssamenstelling van Diu (waarbij de rooms-katholieke gemeenschap met name belangrijk is, want daar vinden we de moedertaalsprekers), als op patronen van meertaligheid onder de sprekers van het Indo-Portugees van Diu. Dit hoofdstuk bevat daarnaast een kort onderzoek naar de geobserveerde variatie in het Indo-Portugees van het eiland, alsmede een beschouwing over de huidige en toekomstige vitaliteit ervan. Hoofdstuk 3 behandelt de sociale geschiedenis van Diu om de vorming en verspreiding van het Indo-Portugees aldaar te verklaren. Het begint met een korte politieke en sociale geschiedenis van het gebied in het bijzonder en de Portugese bezittingen in India in het algemeen. Hierna biedt het een scenario voor het ontstaan van het Indo-Portugees op Diu.

Deel II, 'Beschrijving' ['Description'], bestaat uit een synchrone taalkundige beschrijving van het Indo-Portugees van Diu. Hoofdstuk 4 gaat in op het noodzakelijke descriptieve voorwerk: de gebruikte methoden bij het verzamelen en analyseren van data, de samenstelling van het corpus en de relevante conventies op het gebied van fonemische en fonetische

transcriptie, orthografie en glossering. De taalkundige beschrijving zelf wordt gegeven in hoofdstuk 5 tot en met 8. Hoofdstuk 5 beschrijft de fonologie van het Indo-Portugees van Diu. Hoofdstuk 6 definieert de relevante *Parts of Speech* (lexicale categorieën) en beschrijft tegelijkertijd de morfologische operaties die erin mogelijk zijn (hierbij wordt niet ingegaan op de derivationele morfologie). Hoofdstuk 7 is een overzicht van de syntaxis van de taal, daarbij wordt vooral ingezoomd op de meest in het oog springende eigenschappen, maar ook op de wisselwerking tussen de syntaxis en uitdrukkingen van pragmatische waardes. Hoofdstuk 8 biedt een uitvoerige beschrijving van de lexicon van het Indo-Portugees op Diu (inclusief de derivationele morfologie) en bediscussieert aspecten die relevant zijn voor een comparatief onderzoek en voor de correcte beschrijving van het ontstaan van deze taal.

Deel III, 'Discussie' ['Discussion'], is een vergelijkende studie die enerzijds verder bouwt op de data van Deel II, en anderzijds informatie gebruikt over de talen die het meest van belang waren bij het ontstaan van het Indo-Portugees van Diu. Het laat helder de gevarieerde herkomst zien van typische eigenschappen van de taal en maakt daarmee tegelijkertijd de complexiteit van het onstaans-/ontwikkelingsproces van het Indo-Portugees op Diu inzichtelijk.

334

Curriculum Vitae

Hugo Canelas Cardoso was born in Coimbra (Portugal) in 1980. He graduated in Portuguese and English Studies from the Universidade de Coimbra in 2002, a degree which in effect encompassed both literary studies and linguistics. He finally decided to pursue the linguistics side of his formation, and then went on to obtain an MPhil in Linguistics from the University of Amsterdam in 2003, with a dissertation exploring the Portuguese contribution to the lexicon of Saramaccan. After working as an English-language lecturer at the Instituto Superior Politécnico de Viseu (Portugal) for a year, he began his doctoral research at the Amsterdam Centre for Language and Communication (ACLC) in 2004, which resulted in the present work.